第一章

导　论

本章的目的在于使读者在学习博弈论之前对博弈论有一个总体上的认识，这种认识对于即将开始博弈论学习之旅的读者而言无疑是有益的，它就好比一张地图，为即将扬帆远洋的航船指明了方向和位置。考虑到本书的读者主要是在校的经济学专业本科生，他们接受了将近一年的经济学基础理论学习，但对博弈论却知之不多，有此导论无疑能起到承上启下的作用。这一章主要讨论四个问题，即何谓博弈论、博弈论为经济学带来了什么、博弈论的要点以及本书的主要阅读对象、篇章结构和一些技术上的说明。

VISIT…

LANZAROTE Caliente.com

第一节　什么是博弈论

在回答什么是博弈论之前，我们首先需要明白什么是博弈。按照现代汉语词典的解释，"博"是丰富多彩的意思，而"弈"则指下棋、打牌等对抗性的游戏，因而"博弈"就是指丰富多彩的对抗性游戏。无独有偶，在英语中，"博弈"一词就是游戏(Game)。看来博弈与游戏结下不解之缘绝非偶然，下棋、打牌这类游戏的一个共同特征是所有游戏的参与者都试图"智胜"对手，要智胜对手就需要随机应变，最为重要的是能够预见到对手可能采取的行动，事先有所准备。用专业一点的话来说，就是游戏参与者彼此的策略是相互依存的。为了说明这一点，我们用石头、剪刀、布这个游戏为例。

例 1.1　石头、剪刀、布。

孙悟空和猪八戒护送唐僧去西天取经，一日来到火焰山，方圆百里酷热难当，孙悟空和猪八戒二人必须有一人去寻找淡水，否则唐僧师徒将命丧火焰山。但唐僧师徒经过长途跋涉已非常劳累，在火焰山找水更是难上加难，因而寻找淡水是份苦差事，孙悟空和猪八戒谁都不想承担这份工作。唐僧命令谁去都会引来"凭什么让我去"这样的诘问。一个较好的解决办法就是让唐僧当裁判，孙悟空和猪八戒通过石头、剪刀、布来决定胜负，谁输谁去寻找淡水。由于双方赢的概率一样，因而无论谁输，都输得心服口服(不存在耍赖这种情况)。游戏规则如下：孙悟空和猪八戒必须同时出招，谁后出算谁输；石头胜剪刀，剪刀胜布，布胜石头；如果一样，不算输也不算赢，重新再来一次。例如，孙悟空出剪刀，猪八戒出布，孙悟空赢。如果孙悟空和猪八戒都出剪刀，则重新开始游戏。根据上面所给出的信息，可以把孙悟空和猪八戒两人可能的策略和输赢列成矩阵形式，如图 1.1 所示。

		猪八戒		
		石头	剪刀	布
	石头	未定，未定	休息，找水	找水，休息
孙悟空	剪刀	找水，休息	未定，未定	休息，找水
	布	休息，找水	找水，休息	未定，未定

图 1.1 石头、剪刀、布

从这个游戏可以看出，无论是孙悟空还是猪八戒，他们出什么招，关键在于他们对对手可能出什么招的“猜测”上。如果孙悟空“猜测”猪八戒可能出石头，孙悟空的最优策略是出布，其他招都属不智；如果孙悟空“猜测”猪八戒可能出剪刀，孙悟空的最优策略是出石头，其他招都属不智；如果孙悟空“猜测”猪八戒可能出布，孙悟空的最优策略是出剪刀，其他招都属不智。同理，对于猪八戒而言也是一样。因而我们看到，孙悟空采取什么策略，关键取决于猪八戒的策略，而猪八戒采取什么策略反过来又取决于孙悟空的策略。对这种情况正规的表述是孙悟空和猪八戒的策略具有相互依存性。如果我们引入策略函数的概念，策略的相互依存性就表现得更为明显。假设 fs 为孙悟空的策略函数，fz 为猪八戒的策略函数；a_{11}表示孙悟空的策略——石头，a_{12}＝剪刀，a_{13}＝布；a_{21}表示猪八戒的策略——石头，a_{22}＝剪刀，a_{23}＝布。这里我们默认 a 的第一个下标表示是第几个参与者，1 表示孙悟空，2 表示猪八戒；a 的第二个下标表示是第几个策略，1 表示石头，2 表示剪刀，3 表示布。由此可得：

孙悟空的策略函数：$fs(a_{21})=a_{13}$，$fs(a_{22})=a_{11}$，$fs(a_{23})=a_{12}$；

猪八戒的策略函数：$fz(a_{11})=a_{23}$，$fz(a_{12})=a_{21}$，$fz(a_{13})=a_{22}$。

孙悟空的策略函数其自变量是猪八戒的策略，因变量是孙悟空的策略，例如 fs(石头)＝布，即如果猪八戒出石头，孙悟空的最优策略出布。对于猪八戒而言，其策略函数的自变量为孙悟空的策略，因变量是自己的策略。这是典型的策略相互依存性。

这种策略依存现象广泛存在于人类社会的方方面面，最极端的情景就是以命相搏的战争。这就是为什么博弈论思想的萌芽最早产生于

战争,并最早运用于战争。从这点(策略依存)来看,战争、阶级冲突、政治斗争、经济竞争,甚至生物界的优胜劣汰都与游戏并无二致。这就是为什么博弈论不仅被运用于经济学的研究,而且被广泛运用于其他社会科学和自然科学。例如,博弈论被广泛运用于军事、政治、法律等社会领域;在自然科学中,则被运用于人工智能、物种演化等领域。为了加深理解我们再举几个有趣的例子。

例 1.2　诺曼底登陆。

诺曼底登陆是二次大战中的重要战役,它对加速法西斯德国的垮台具有重要的战略意义。当时盟军在法国登陆的地点有两个,一是马赛,二是诺曼底。由于盟军的军力有限,不可能同时在两个地点登陆。由于相同的原因,德军也不可能同时在两地设防,只能集中兵力在一地设防。如果盟军在马赛登陆,而德军正好在马赛设防,那么由于德军以逸待劳,再加上天时地利,将能够击败盟军的进攻。相反,如果盟军在马赛登陆,而德军在诺曼底设防,盟军将不会遇到有效的抵抗,登陆将会成功。我们可以将历史上的诺曼底登陆之战抽象为图 1.2。

		德军	
		马赛设防	诺曼底设防
盟军	马赛登陆	失败,成功	成功,失败
	诺曼底登陆	成功,失败	失败,成功

图 1.2　诺曼底登陆

如果盟军在诺曼底登陆,而德军在马赛设防,那么盟军登陆成功而德军溃败,其他同理。从图 1.2 可以看出,盟军之所以在诺曼底登陆,是因为盟军相信德军会在马赛设防,而德军之所以在马赛设防是因为它相信盟军会在马赛登陆。也就是说,盟军的最优策略是德军策略的函数;反之,德军的最优策略是盟军策略的函数,即盟军和德军的策略是相互依存的。

例 1.3　鸽派和鹰派。

前苏联和美国是冷战时期的两个超级军事大国,并长期处于对抗状态。从历史来看,前苏联和美国很少出现公开的直接冲突,通常是美

国强硬，苏联就会采取暂时回避的策略，如果是苏联强硬，美国就会主动回避。像古巴导弹危机那样的紧张局面极少出现，即使出现双方都会采取妥协的策略。前苏联和美国之间的这种现象我们可以用一个简单的博弈模型来加以概括。显然，如果前苏联和美国之间发生直接冲突，后果有可能是同归于尽，因而妥协要比一味强硬结果要好。在政治术语中，鸽派通常代表妥协路线，而鹰派通常代表强硬路线。前苏联和美国之间的博弈如图 1.3 所示。

		美国：鸽派政策	美国：鹰派政策
前苏联	鸽派政策	0,0	−1,+1
前苏联	鹰派政策	+1,−1	−ω,−ω

图 1.3 鸽派和鹰派

图 1.3 表明，如果前苏联和美国都采取鸽派政策（避免冲突），那么双方得到的利益为零，表示双方相安无事。如果双方都采取强硬路线，那么必然导致冲突的发生，后果是双方同归于尽，得到的利益为负的无穷大。如果一方采取强硬路线，而另一方采取鸽派政策，那么实行鹰派政策的一方可以威吓对方从而获得较多的利益，用+1 表示，而实行鸽派政策的一方则失去一部分利益，用−1 表示。因此，前苏联和美国到底采取什么政策，主要取决于彼此认为对方可能采取的策略。如果美国采取鸽派政策，那么前苏联的最优反应就是表现出强硬立场。如果美国采取鹰派政策，那么前苏联的最优策略就是退避三舍，采取鸽派政策。上述这个模型虽然简单，但却真实地反映了冷战时期的本质特征。在图 1.3 中，存在着两个博弈均衡解，它们分别是（前苏联鸽派政策，美国鹰派政策）和（前苏联鹰派政策，美国鸽派政策）。这就为冷战时期美苏两个超级大国主动避免冲突的行为提供了有说服力的解释。

通过上面的说明，现在我们给出博弈的正式定义。

定义 1.1 博弈是指决策主体（个人、企业、集团、政党、国家等）在相互对抗中，对抗双方（或多方）相互依存的一系列策略和行动的过程集合。

在这个定义中有几点要注意：

第一，博弈中的参与者各自追求的利益具有冲突性。如果决策主体之间的利益是一致的，就不是博弈。从学术观点来看，即使一个博弈包含无穷多参与者，因利益一致，也可以理解为一个人。由于一个人是不会和自己博弈的，因此从某种意义上看，博弈论是一门研究冲突的学科，它为人们理解冲突和合作提供了一种重要的思想方法。

第二，博弈是一个过程集合。博弈不是一个孤立的事件，而是人们在对抗过程中有关的所有方面的集合。它包含参与者的集合、策略的集合、行动的集合、信息的集合等。把博弈看作一个集合是思维从具体到抽象的重要一步。

第三，博弈的一个本质特征就是策略的相互依存性。如果博弈参与者之间的策略不存在依存性，那么与一个人自娱自乐的游戏并无区别，当然也就不能称其为博弈。不过在一种特殊的境况下，有一种博弈“不存在”策略的相互依存性，这种博弈就是包含严格优策略的博弈。

例 1.4　如图 1.4 所示的囚徒困境。

		囚徒 2	
		沉默	招供
囚徒 1	沉默	−1，−1	−9，0
	招供	0，−9	−6，−6

图 1.4　囚徒困境

警方抓住了两个犯罪嫌疑人——囚徒 1 和囚徒 2，但是警方掌握的证据并不充分，因而需要犯罪嫌疑人的口供。警方将两个囚徒分别关在不同的房间，但告诉两个囚徒相同的信息，即如果一方沉默而另一方招供，那么招供方立即释放，沉默方将被判入狱 9 个月；如果双方都沉默，那么两人将在 1 个月后因证据不足而释放；如果双方都招供，那么两人都将被判入狱 6 个月。在这种情景下，对于任何一个囚徒而言，无论对方选择什么策略他的最优策略都是招供，因而出现了囚徒 1 不“依赖”囚徒 2 的策略来选择最优策略的情况（反过来也一样）。招供就是两个囚徒的严格优策略。

明确了什么是博弈，对博弈论就好下定义了。

定义 1.2 博弈论是专门研究博弈如何出现均衡的规律的学科。

正是由于均衡在博弈论中具有中心地位（均衡在经济学中也同样如此），因而博弈在经济学中运用最广，并主要由经济学家加以发展。1994 年诺贝尔经济学奖授予了对博弈论发展作出贡献的三位学者，其中一位是数学家、两位是经济学家。正是由于均衡的重要性，我们将 Game Theory 译作博弈论而不是对策论，因为对策论是一个数学命题，而均衡是一个经济学命题。按照中国大百科全书的解释，均衡是“经济学中，用以分析考察经济中若干变量之间的关系，以解释经济现象及其变化的概念或分析方法”。因而博弈论体现的是一种经济学思想，不过这种思想不仅仅局限于经济学本身。

第二节 博弈论革命

在经济学界还没有人把博弈论的发展对经济学的影响称为革命，但客观看来，博弈论带给经济学的东西不仅仅是一个“革命”所能概括的，尽管这种影响是在一种潜移默化并且漫长的过程中完成的。它平平淡淡，但远胜轰轰烈烈。自 20 世纪 90 年代以来，有三次诺贝尔经济学奖授予了与博弈论相关的研究领域（1994 年的博弈论，1996 年和 2001 年的信息经济学），总共有八位经济学家因此而荣获诺贝尔经济学奖，是获奖人数最多的经济学研究领域（排第二位的是金融领域）。窥一斑而知全豹，再过十年，博弈论对经济学的影响会表现得更为明显。

经济学研究离不开四个重要范畴：决策主体、行为、制度结构和均衡。

决策主体，又称经济行为主体（Actor）：经济行为主体是经济生活中追求自身利益最大化的决策者和参与者。例如消费者、生产者、政府以及任何的利益集团都可以在经济学中被看作是决策主体。

行为（Behavior）：是经济行为主体在面对各种约束条件下的最大

化行为。最典型的例子就是消费者在收入、市场价格等的约束下使自身的效用最大化；生产者在成本、技术、市场价格等的约束下，使利润最大化。

制度结构(Institutional Structure)：制度结构可以看作是经济行为主体面对的所有约束，它规定了经济行为主体拥有什么条件去选择，如何去选择，能选择些什么，同时它也受到行为主体自身选择和其他行为主体的选择，以及彼此互动关系的决定。

均衡(Equilibrium)：均衡就是一个能够得以维持的结果，或者说是所有行为主体不得不接受(可能愿意也可能不愿意)而又不可能更好的结果。

以上四个范畴构成了经济学研究的基石。然而，在传统的经济学研究中，通常主要考察的是行为主体如何选择，最终如何得到均衡，如何把制度结构作为上天给定的外部条件来加以处理的。改变不同的制度结构(实际上在经济学中这相当于更改约束)，就得到不同的均衡结果。但是对于制度结构本身的性质和演变，传统经济学不将它作为考察的对象，同时也不具有这样的能力。对经济学而言，排除对制度结构的考察是一个巨大的遗憾，它不仅大大制约了经济学的研究领域，而且经济学的有效性和科学性也受到了广泛的质疑。因而考察制度结构可以说是现代经济学和传统经济学的一个根本区别。

一般均衡理论是整个经济学的理论基石和道义基础。一般均衡理论表明市场机制是完美的。这一点集中体现在福利经济学第一定理和第二定理上。在完全竞争的市场和民主政府下，平等与效率可以兼顾。然而在下列情况下，一般均衡理论是不成立的：

1. 非完全竞争。在现实的经济生活中，特别是进入20世纪后，垄断越来越成为一种普遍而占主体的经济现象和制度结构。在垄断下，完全竞争的效率不复存在。

2. 外在性。由于行为主体之间客观存在着密切的联系，因而个体并非是孤立地存在着，一个行为主体的行为有可能对其他人造成影响，这种现象就是外在性，如污染。外在性的存在表明市场不可能把所有的成本收益都计算在内，有些成本和收益客观地存在着，但市场却无能力

对此进行计算。这说明市场是有缺陷的。

3. 公共产品。在现实生活中,公共产品无处不在,可以说它和私有物品同样普遍,但是在一般均衡理论中,公共产品是不考虑的。其原因是在公共产品的消费上存在着"搭便车"的行为,而市场机制本身无法解决"搭便车"问题。

4. 逆向选择问题。随着经济学研究的深入,人们发现信息不对称同样会造成市场的失灵,其中最突出的是逆向选择和道德风险问题。2001 年诺贝尔经济学奖的三位得主就是因在这两个领域作出了开创性的工作而获奖。阿克洛夫(Akerlof)提出的柠檬(Lemon)问题清楚地表明市场不仅不能解决这类问题,相反在逆向选择下会彻底消失。针对这种情况,斯彭思(Spence)提出可以通过向行为主体发信号来解决逆向选择问题。现今,信号博弈已发展成为信息经济学的两大模型之一。

5. 道德风险问题。科斯(Coase)的企业理论被认为是新制度经济学的奠基之作。科斯认为之所以产生企业,是由于市场与企业(制度)相比交易费用太高,因而出现了企业,并替代市场。但是对于为什么市场交易费用高、交易费用的性质是什么等根本性的问题,科斯的回答却是差强人意的。在随后的研究中,人们才逐步发现这些问题不过是道德风险问题的一部分。因而如何解决道德风险问题成了新制度经济学和企业理论的核心课题,并在此基础上提出了各种各样的委托一代理模型。

对于上面所谈的市场失灵的五个方面,传统的经济学缺乏有效的手段去加以研究。由于一般均衡是在排除这五个方面后得出的结论,因而当面临这五个问题时,一般均衡就有可能是不存在的。这构成了经济学的致命伤。一个世纪以来,经济学的主要发展实际上都是围绕着这五个问题展开的,例如垄断竞争理论、产业组织理论、企业理论、信息经济学、新制度经济学、不确定下的决策(投资理论)以及宏观经济学。

博弈论的发展与成熟实际上使得经济学在面对上述五个问题时,有了作出深入研究的基础,而逆向选择和道德风险问题本身就是博弈论在发展过程中提出的。此外,博弈论使得经济学研究制度结构的来龙去脉及其演变成为可能。对于传统经济学而言,制度结构是作为外生变量给定的,因而经济学实际上几乎成了专门研究给定条件下如何规划

利益最大化问题的学科，而这实际上是数学就能胜任的事。经济学日益丧失它的社会性，必然导致经济学的贫乏和失去其存在的必要。从这个意义上看，博弈论对于经济学而言是一场有着深远影响的革命。它极大地丰富了经济学的内容，使得经济学的研究领域和经济学的适用范围大大扩展；也使得经济学家的思维从一种静态、孤立的模式发展成互动、联系的模式，从而彻底改变了经济学家在观察问题和解决问题时的思维方式。这一点连马克思主义学者都加以承认。由于马克思关心现代社会的矛盾及其解决方法，因此博弈论简直就是为马克思主义的进一步发展所“定做”(Tailor-made)的工具。①

博弈论革命不仅表现在经济学中，也体现在其他社会科学中。

萨缪尔森(Samuelson)在其著作《经济学》中曾指出，在社会科学研究中人们常常会犯“合成谬误”(Fallacy of Composition)式的推理错误，即认为个体理性必然意味着集体理性。这一推理的错误性通过下面的例子可以看出，当许多人坐着看表演时，坐在后面的人站起来可以看得更清楚，因而站起来对个人而言是一种理性行为，但当所有人都站起来后，除了第一排的人看得清楚以外，其他人的境况都会变得更差。又如，股市上个人抛出股票实属正常，但当所有人都抛出股票时，面对的可能就是股市崩溃。对于个人而言，“随大流”显然是理性行为，但个体理性导致的却是整体非理性。前面提到的囚徒困境一例，更反映出个体理性与集体理性是冲突的(个体理性要求都招供，而集体理性要求都沉默)。

正是基于这些观察，美国著名的政治学家奥尔森(Mancur Olson)于1965年发表了经典性著作《集体行动的逻辑》，② 探讨了利益集团集体行动的微观基础。在奥尔森看来，利益集团的集体行动的目标是给每个成员提供公共产品。但公共产品具有非独占性和非排他性，因而存在着“搭便车”的行为。这意味着在利益集团内部，集体行动的目标和个体

① Jon Elster, 1982:"Marxism, Functionalism and Game Theory", in *Theory and Society II*, pp. 453～482.

② 奥尔森，《集体行动的逻辑》，上海三联书店，1990年。

理性存在着矛盾，集团并不像一个人那样有理性，而集体行动能否成功实际上又取决于集团内每一个成员的努力。因而奥尔森从单个成员对集团的公共产品的贡献与他从集体产品中得到的份额的关系入手，以此来对不同集团的内部结构进行分类。奥尔森描述"集体行动的逻辑"的最简单的方程为：

$$P_i = R_i - C$$

其中，C 是获取公共产品的总成本，R_i 是第 i 名成员从总的公共产品中得到的总收益，P_i 是第 i 名成员的净收益。若对某些(少数)i 而言，$P_i > 0$，则奥尔森将这样的集团定义为"含有特殊人物的集团"；若对所有 i 而言，$P_i < 0$，则奥尔森将这类集团定义为"隐集团"。

在"含有特殊人物的集团"中，这些特殊人物从公共产品中得到收益或好处是如此之大，以至于超过获取公共产品的总成本，因此他们愿意负担全部成本以获得公共产品。如果一个集团中含有这样的特殊人物，那么该集团在获得公共产品方而就能够成功。例如，在股份公司中，大股东就是利益集团(全体股东)中的特殊人物，通过大股东对经理实行监督，小股东则"搭便车"得到丰厚的回报。下面的例子说明了这种情况。

例 1.5　富人与穷人。

		穷人：巡逻	穷人：不巡逻
富人	巡逻	6,0	4,2
	不巡逻	8,0	0,0

图 1.5　富人与穷人博弈

在一个小区内，住着一个富人和一个穷人。组织夜间巡逻能够有效防止偷盗，但夜间巡逻的成本为 4。假设富人的财产为 8，穷人的财产为 2。如果两人都巡逻，那么巡逻成本由两人均摊；如果只有一人巡逻，则由巡逻者承担。如果富人和穷人都巡逻，那么富人的收益为 8－2＝6(财产－成本＝收益)，而穷人的收益为 2－2＝0；如果富人巡逻而穷人不巡逻，那么富人的收益为 8－4＝4，穷人"搭便车"，财产得保，收益为

2－0＝2；如果两人都不巡逻，财产被窃贼偷光，收益都为零。从图1.5可以看出，如果富人巡逻，那么穷人最优的策略是“搭富人的便车”不巡逻。问题是，在穷人不巡逻的情况下，富人会不会巡逻，答案是肯定的，因为巡逻有4的收益，不巡逻收益为0。因此该博弈的均衡解是（富人巡逻，穷人不巡逻）。富人就是利益集团（住宅小区居民）中的特殊人物。

在“隐集团”中，奥尔森又区分出“大集团”和“中等集团”。此时，由于$P_i<0$，没有特别人物自愿承担集体产品的全部成本，每个成员从个人理性出发，将不会自愿分担集体产品的总成本。其原因是他们自以为：

1. 由于集体产品的非排他性，不付出成本也可以享受；

2. 自己付出努力，但并不能保证别人也会分担公共产品的成本，即存在着“搭便车”的现象；

3. 每个人对公共产品的贡献都很小，无关大局。

据此，奥尔森得出了“集体行动的逻辑”的主要结论：小集团要比大集团有力量，因为搭便车的现象少。对于“隐集团”来说，如果没有外部强制或对成员的特别激励，将不会有提供公共产品的集体行动。

通过上面的介绍我们可以看到，奥尔森所谓的“集体行动的逻辑”实际上与囚徒困境的逻辑完全一样。1971年，哈丁（Hardin）在其论文中，证明了“集体行动的逻辑”与n人囚徒困境实际上是等价的。①

罗尔斯（John Rawls）的《正义论》是20世纪下半叶人类思想史上的一部重要著作。自1971年出版以来，受到了伦理学家、政治学家、经济学家、社会学家和法学家的极大关注，此书的影响深度在20世纪下半叶难有出其右者。

在《正义论》中，罗尔斯试图证明在“原始状态”的“无知之幕”的条件下，人们如何通过理性的选择得到正义原则，以在一组“社会的基本结构”上配置各个社会成员的“基本益处”。“原始状态”下的“无知之幕”是指每一个人都不知道自己的出身、社会地位、财富、能力、智慧和

① R. Hardin, 1971:“Collective Action as An Agreeable n-Prisoner's Dilemma”, *Behavioral Science*, No. 5, pp. 472～479.

爱好，从而保证没有人能够在选择和设计正义原则时塞进“私货”。在罗尔斯看来“基本益处”是每一个人都想要的东西。它包括基本的自由、机会、收入、财富和权力，以及（罗尔斯认为特别重要的）自尊。要配置“基本益处”需要一系列社会的、政治的、法律的和经济的制度。这个制度集合就是社会的基本结构。例如，宪法、所有制、市场机制、一夫一妻制等。罗尔斯在推导正义原则时，假设每一个人合理的选择是遵循“极小极大”的决策原则，在“极小极大”原则下，人们将发现罗尔斯所提出的正义原则要优于其他的正义原则（例如功利主义原则）。

从根本上说，社会契约是非零和博弈的合作解，[①]而罗尔斯旨在“将洛克、卢梭和康德所代表的传统社会契约论一般化和达到更高度的抽象”，[②]因此罗尔斯的分析实际上完全等同于博弈论中的动态博弈。又由于“极小极大解”只不过是博弈论中最原始的均衡解概念之一，因而罗尔斯的分析又是远远不够的。博弈论的发展不仅指出了罗尔斯《正义论》的不足，同时也指明了继续前进的方向。

在自然科学领域，博弈论同样有着广阔的运用前景，如人工智能、生物种群的演化等博弈论都显示出巨大的威力。1997 年计算机“深蓝”战胜被称为棋王的国际象棋大师加里·卡斯帕罗夫，曾引起世人的巨大关注。深蓝的胜利与其说是机器的胜利，不如说是人类思想的胜利，因为“深蓝”下象棋的程序就是按照博弈论中的动态博弈的思想来编程的。

综上所述，把博弈论的发展对人类思想的贡献称之为“革命”并不为过，至少对经济学是如此。

第三节 博弈论要点

在这一节里，我们对博弈作一个总体但是简单的介绍，正式的叙述

① M. Taylor, 1976: *Anarchy and Cooperation*. Wiley.

② J. Rawls, 1971: *A Theory of Justice*, VIII.

会在专门的章节中展开。

一、博弈的四个要素

博弈的形式多种多样，理论的力量就在于它能从各种具体的形式中抽象出一般。通过前面的例子可以悟出，任何一个博弈必然包含四个要素，这四个要素也是从理论上描述一个博弈不可或缺的建筑材料。

(1)参与者(Player)：参与博弈的决策主体。判断博弈参与者的根本标志是他(或她)是否是博弈的利害关系者，只有在博弈中存在利害关系的决策主体才被看作是博弈的参与者。一个较为特殊的、与其他参与者不存在利害关系的参与者我们把它定义为自然(Nature)。

(2)博弈的规则(Rule)：对博弈作出具体规定的集合。它包含对参与者行动顺序的规定、当某个参与者行动时他所知道的信息、有什么样的行动可供选择、选择之后会得到什么样的结果，等等。

(3)结果(Outcome)：对所有参与者的每一个可能的行动组合，会出现什么样的结果。

(4)收益(Payoff)：在可能的每一个结果上，参与者的所得和所失。用经济学的话来说，就是在所有可能的结果上参与者的偏好是什么。这意味着，博弈的每一位参与者会在不同的结果之间进行比较，以寻求最好的收益。

一般而言，有什么样的结果就会有什么样的收益，因而在分析博弈的时候可以把结果和收益看作是一回事(这在分析上有好处)。不过严格来说，收益是结果的一个函数(pay off $=f$(out come))。例如囚徒困境，如果囚徒 1 招认，囚徒 2 沉默；结果是囚徒 1 释放，囚徒 2 入狱 9 个月；囚徒 1 的收益为 0，囚徒 2 的收益为 -9。

例，石头、剪刀、布。

(1)参与者：孙悟空和猪八戒。通常用阿拉伯数字来代表，将孙悟空定义为 1，猪八戒定义为 2。

(2)博弈规则：双方同时出招，不得耍赖；石头赢剪刀，剪刀赢布，布赢石头。

(3)结果：有两种可能，(孙悟空休息，猪八戒找水)或是(孙悟空找

水，猪八戒休息）。

(4)收益：孙悟空休息即得到正的效用，比如为+U；孙悟空找水得到负的效用，比如为-U。所以孙悟空偏好休息甚于找水，即休息总是比找水要好。

例，诺曼底登陆。

(1)参与者：盟军和德军，分别定义为1和2。

(2)博弈规则：盟军和德军彼此不知道对方的计划，如果两军相遇，盟军惨败；如果不相遇，盟军胜利。

(3)结果：有两种可能，(盟军胜利，德军惨败)或是(盟军惨败，德军胜利)。

(4)收益：盟军胜利得到的收益为$+\infty$，德军惨败得到的收益为$-\infty$；盟军惨败得到的收益为$-\infty$，德军胜利得到的收益为$+\infty$。因而盟军和德军都希望胜利而不是惨败。

例，烛光晚餐。

假设有一对情侣，各自在不同的地点工作。下班后，双方由于客观原因无法联系上，但是他们希望能够在一起共进晚餐。他们常去吃饭的地点有两处，一是利顺德，二是喜来登。如果他们能碰在一起，双方共进烛光晚餐，感到非常幸福；如果遇不到一起，将各自在闷闷不乐中进餐。

(1)参与者：情侣1和情侣2。

(2)博弈规则：两个参与者分开，并且不能通信。他们必须各自作出在何处就餐的决定。

(3)结果：要么共进晚餐，要么黯然神伤。

(4)收益：共进晚餐得到的效用为+U，分开进餐得到的效用为-U。因而情侣1和情侣2偏好共进晚餐胜于黯然神伤。

任何一个博弈都具有这四个要素。对这四个要素进行科学的描述和表达，就为博弈的表达式。在博弈论中，博弈的表达式只有两种——基本式和扩展式。下面我们分别加以介绍。

二、博弈的基本式(Normal Form)

如果用参与者、策略和收益函数来表述一个博弈，就称为博弈表达

的基本式。

我们用囚徒困境来说明如何把一个博弈转变为基本式。根据前面的已知条件,写下囚徒困境的四个要素。

(1)参与者:囚徒 1 和囚徒 2。

(2)博弈规则:坦白从宽,抗拒从严。

(3)结果:要么 1 人释放,1 人入狱 9 个月;要么两者都入狱 6 个月;要么两者入狱 1 个月。

(4)收益:释放收益为 0,入狱 9 个月收益为 −9,入狱 6 个月收益为 −6,入狱 1 个月收益为 −1。

将上述条件转变为基本式。

(1)参与者集合:定义为 I={1, 2}。其中,1 表示囚徒 1,2 表示囚徒 2。

(2)策略集合:定义为 S_i={招供,沉默},i=1,2。S_i 又称为策略空间。策略 s_{11}=招供,s_{12}=沉默,s_{21}=招供,s_{22}=沉默。第一个下标表示第几个参与者,第二个下标表示第几个策略,例如 s_{11}=招供,表示囚徒 1 的第一个策略为招供。其他依此类推。

(3)收益函数:给定所有参与者的策略组合,每一个参与者得到的收益。策略组合就是每一个参与者的策略组成的向量。收益函数使得策略组合和收益之间建立了一种函数关系,定义为 $u_i(s_{1j}, s_{2j})$,j=1,2 表示第几个策略。例如,$u_1(s_{11}, s_{21})=-6$,即囚徒 1 招供,囚徒 2 招供,囚徒 1 得到的收益为 −6(6 个月监禁)。当我们考察所有的策略组合时,通过收益函数,相应地得到不同的收益,将它们列成矩阵形式,就得到前述图 1.4 所示。因而在博弈论中,又把博弈的矩阵表示等同于基本式。

为什么我们把包含上述三个要素的表达式称为"基本"式?它"基本"在什么地方?孤立地解释显然不好理解,而结合博弈的扩展式表达就非常容易说明。

三、博弈的扩展式(Extensive Form)

表达博弈的第二种方法就是博弈的扩展式。如果我们详细地描绘

出一个博弈的参与者、参与者的策略、参与者的行动顺序及其行动时所拥有的信息、可能的结果和收益，其细节就称为博弈的扩展式。扩展式之所以称为扩展式就是因为它比基本式"详细"。基本式中许多省略的细节(这就是我们为什么要把它称为基本式的根本原因)，在扩展式中都得到了详细的说明。我们还是以囚徒困境为例。前面博弈矩阵(基本式)表示的囚徒困境我们也可以用博弈树(扩展式)来等价地表示。如图1.6所示。

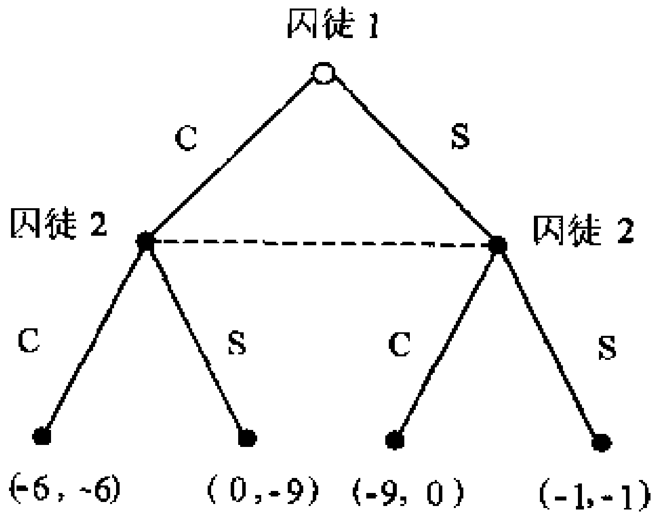


图 1.6 囚徒困境

这里我们不妨假设囚徒1首先行动，囚徒2随后采取行动(也可以假设囚徒2首先行动，然后囚徒1采取行动，这表明基本式转化为扩展式的形式不是惟一的)。在第一个节点标注囚徒1，表示囚徒1第一个行动。第一个节点称为起始点，通常用空心圆表示。它有两个分支，分别代表可采取的两种行动——招供(用C表示)和沉默(用S表示)。在第二节点和第三节点标注囚徒2，表示囚徒1行动后，轮到囚徒2行动。囚徒2也有两种可采取的行动——C和S。之后博弈结束，囚徒1和囚徒2分别得到各自的收益，第一个数值表示囚徒1的收益，第二个数值表示囚徒2的收益。这里我们看到囚徒2所处的两个节点用虚线连接起来，这表明囚徒2在行动的时候，并不清楚囚徒1采取的行动是C还是S，当然囚徒1也不清楚囚徒2采取什么行动(因为他在囚徒2前采取行动)，因而就好比他们俩同时行动(正是这一点使得它与基本

式等价)。打个比方,囚徒1首先被警方抓获,随后囚徒2也被警方抓获,囚徒1和囚徒2不存在串谋的可能。显然,囚徒2不知道囚徒1是否招认,囚徒1同样也不知道囚徒2是否招认。因而这个博弈即可以用图1.4的博弈矩阵——基本式来表示,也可以用图1.6博弈树——扩展式来表示。

显然,博弈树要比博弈矩阵详细得多。它清楚地表明了参与者的行动顺序(虽然这里不重要)、拥有的信息、信息的特征、[①] 策略的本质等。例如,在基本式中,囚徒1和囚徒2的策略和行动是同一的,都为招认和沉默。如果我们将图1.6中连接囚徒2的虚线取消,那么在扩展式中,囚徒1的策略为招认和沉默,囚徒2的策略则变为4个,即(如果囚徒1招认,那么招认)、(如果囚徒1招认,那么沉默)、(如果囚徒1沉默,那么招认)、(如果囚徒1沉默,那么沉默)。扩展式清楚地揭示出了策略的本质,即策略是随情况而定的完整的行动计划。行动并不等同于策略,策略是行动的指南,而行动只不过是策略的实践。在基本式中,囚徒2的策略就是行动,即招认和沉默,并没有体现随囚徒1策略的变化而变化这样的特征。也许我们会产生这样的疑问,是不是博弈的扩展式要比基本式"好"?从理论上来讲,任何基本式都可以转化为扩展式,而任何扩展式也可以转化为基本式。[②] 用哪一个表达式关键是看哪一个方便。通常用基本式来表示静态博弈较为方便,而用扩展式来表示动态博弈则较为方便。对于博弈的基本式和扩展式,我们在后面的章节中还会进行正式而严格的讨论。

四、信息和顺序

首先,在一个博弈中,信息显然是至关重要的,孙子兵法中的"知己知彼,百战不殆"说的就是这个意思。掌握的信息不一样,得到的结果也必然不一样,因为信息是作出正确决策的关键。有些信息对于博弈参与

① 随着学习的深入,我们会看到所有的完全信息静态博弈都是信息不完美的博弈。

② 对于基本式和扩展式是否等价,理论界仍然存在争论。但从直觉来说,越复杂的扩展式,通过越"详细"的定义策略,就可以使得基本式无限逼近扩展式,因而基本式和扩展式是等价的。不过,用哪一个表达式的标准主要是看它能否达到要求和简便。

者而言是共同信息，而有些信息则只有自己知道。如果所有重要的信息都是共同信息(Common Knowledge)，那么我们称这样的博弈为完全信息博弈。如果一些重要的信息不是共同的，只有当事人知道，或者说不是所有的博弈者都知道，这样的博弈我们就称为不完全信息博弈。在博弈论中，对博弈的完全信息和非完全信息有着严格的定义：如果所有参与者在给定任意策略组合下，每一个参与者的收益都是确定的(包括期望值)，那么就是完全信息博弈。如果至少有一个参与者的收益是不确定的(不确定是指参与者主观认为收益具有多种可能性)，[①] 那么就是非完全信息博弈。

上述这个定义等价于每个参与者的收益函数是否为共同信息，如果每个参与者的收益函数为共同信息，那么每一个参与者的收益必然是确定的。收益不确定是由于有些参与者只能"猜"他人或他自己可能的收益，为什么"猜"？因为参与者的收益函数不是共同信息。共同信息也译为共同知识，将它定义为：你知道这件事，他知道这件事，你知道他知道这件事，他知道你知道这件事，你知道他知道你知道这件事，你也知道他知道你知道他知道这件事，等等，以至无穷。通俗地说，就是博弈的任意两个参与者双方彼此心里都清楚地知道对方知道此事。与共同信息相对应的就是私人信息。私人信息就是一部分人知道，或是所有人不知道，但自然知道。非完全信息博弈比完全信息博弈复杂，分析起来也较困难，但非完全信息较符合实际情况，因而适用范围更广，理论意义更高。

其次，博弈参与者行动的顺序也很关键。他们可以同时行动，也可以有先有后。如果博弈双方同时行动(在博弈论中，所谓同时行动是指一方在作出行动时并不清楚对手是否已经作出了行动)，我们称之为静态博弈。如果一方在作出行动时，知道对手已经作出了行动(可能不知道具体行动是什么，就像图 1.6 囚徒困境)，我们称之为动态博弈。例如，游戏——石头、剪刀、布就是一个静态博弈，而下象棋则是一个动态

---

① 如果我们把不确定性处理成服从一定的概率分布，实际上就等于把非完全信息处理成了完全信息。随着学习的深入，读者会逐步明白这一点。

博弈。从这里可以看出所谓动态与静态的划分与时间先后无关，而与博弈的规则有关。

从上面这两点出发，我们就可以把博弈分为四类，如表 1.1 所示。

表 1.1 博弈的分类

	完全信息	非完全信息
静态	完全信息静态博弈	非完全信息静态博弈
动态	完全信息动态博弈	非完全信息动态博弈

四个博弈之间的关系可用图 1.7 来描绘。

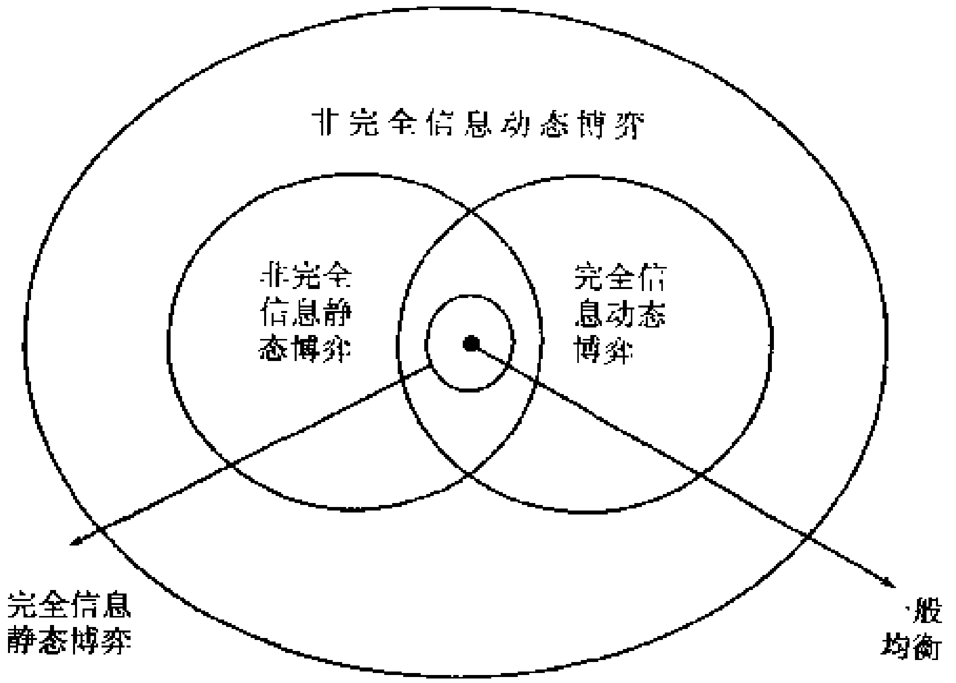


图 1.7 非完全信息的动态博弈等于整个世界

图 1.7 清楚地说明了四个博弈模型之间的关系，最简单的模型为完全信息静态博弈模型，但它对应的现实状况也是最简单的。非完全信息动态博弈模型是最为复杂的模型，因而它对应着最为复杂的现实状态，其他的模型都只不过是它的特例而已。正是因为非完全信息动态博弈几乎与所有现实对应，所以至今非完全信息动态博弈仍然没有一个一般性的解。

从图 1.7 我们还可以看到，一般均衡只不过是非完全信息静态博

弈的一个点而已。如果把一般均衡比喻为鲁宾逊荒岛，那么非完全信息动态博弈对应的则是整个世界，博弈论就像一艘航船，使得鲁宾逊可以离开荒岛探索世界。

第四节 本书的阅读对象

本书的阅读对象为高等院校经济学专业高年级本科生或研究生一年级，他们在学习微观经济学时，已经接触到一些博弈论的基础知识，但并不深入也不成体系。在这本书中，作者努力做到对博弈论有一系统的说明，同时尽力做到深入浅出，重在使学生理解并能加以运用。为了能够很好地理解博弈论，要求学生具有一定的经济学知识和数学知识。虽然作者会努力尽可能详尽地加以解释和说明，但思考还是读者自己的事。如果读者遇到不懂的地方，除了“好问”之外，多动笔也十分关键。“好记性不如烂笔头”，动笔自有一番新天地。最后，学好这门课还需要反复多看几遍，有机会的话阅读书后的参考书那就更好。

本书的结构分为两部分四章。第一部分为静态博弈，它由两章组成：第二章——完全信息静态博弈和第三章——非完全信息静态博弈。第二部分为动态博弈，它也由两章组成：第四章——完全信息动态博弈和第五章——非完全信息动态博弈。在书中用楷体字的部分，主要是一些证明和扩展，读者跳过并不影响阅读。

学习是一件痛并快乐着的事，如果你能读完本书，相信它能对你今后的人生有所助益。最后，祝大家一路顺风，到达最后一页！

复习思考题

1. 什么是博弈论？为什么说博弈论是一场“革命”？你同意这样的观点吗？

2. 请你根据博弈的定义，看一看你身边哪些情况可以描述成博

弈，请举出具体的例子，并将它们抽象成博弈的四个要素。其中这些现象哪些是静态的？哪些是动态的？哪些是信息完全？哪些是信息不完全？

3. 你认为博弈论对你有些什么帮助？是在工具上还是在思想上？

第一部分 静态博弈

在这一部分，我们主要分析静态博弈。所谓静态就是博弈的参与者在行动时，彼此不知道对方的行动是什么。显然，如果参与者“同时行动”，那么参与者彼此必然不知道对方的行动是什么。如果知道了，例如石头、剪刀、布，那么后行动的人必定总是赢，先行动者必定总是输。显然这样的博弈没有什么理论价值。正是因为“同时行动”具有这种特征，博弈论中就把所有参与者“同时行动”的博弈称为静态博弈。这里要注意的是，同时行动和时间上的同时并不是一回事，尽管它可能在时间上是同时的。

在完全信息下的静态博弈中，纳什均衡是核心，抓住了纳什均衡这个纲，就能纲举目张。在非完全信息下的静态博弈中，产生了一个新的纳什均衡概念——贝叶斯纳什均衡，它实际上是非完全信息下（不确定条件下）的纳什均衡。两个模型中求解均衡解的基本方法都是解极值问题。

第二章

完全信息静态博弈

"纳什均衡"这个概念正是在完全信息静态博弈这个基础上产生的,完全信息静态博弈是博弈论中最简单的博弈模型,但却是博弈论产生的基础。纳什(Nash)在1950年发表的不足千字的论文[1]中,首次在思想上证明了完全信息静态博弈存在着一般的均衡解,这个一般均衡解不过是博弈数学模型中的不动点。正是有了"纳什均衡"这个概念,博弈论的研究才真正走上了正确的方向。为了纪念纳什的贡献,后人把这个均衡概念称为纳什均衡。

什么是完全信息?是不是所有当事人都无所不知?还是所有信息都是公共信息?显

① Nash, J. 1950,"Equilibrium Points in n-Person Games", *Proceeding of the National Academy of Sciences* 36; pp. 48~49.

然要作出以上这些判断是非常困难的，因为不同的情景要求的信息是不一样的。这就要求对完全信息要给出一个本质上的定义。在博弈论中对完全信息的定义非常的简单但严格，即如果所有参与者对他(她)们的收益的信念都是确定的，那么这个博弈就是完全信息的。也就是说，所有参与者的收益函数是公共信息。只要有一个参与者的收益函数不是公共信息，那么这个博弈就属于非完全信息。例如在石头、剪刀、布这个博弈中，由于孙悟空和猪八戒的收益函数是公共信息，所以当猪八戒出策略 x(例如剪刀)，孙悟空出策略 y(例如布)，猪八戒和孙悟空就能计算出他们各自会得到多少收益(例如＋U 和－U)。这里大家要注意的一点是，收益函数是公共信息等于说每一个参与者都能明确地计算出彼此的收益，因此说收益函数是公共信息并且是确定的可能更精确一些。

在这一章中，我们主要介绍博弈表达的基本式，并使读者逐步掌握如何将一个博弈转换为基本式，并从简到难地给出纳什均衡这一重要概念。最终的目的是使读者能够运用所学的知识解出各种各样的完全信息静态博弈的纳什均衡解。

第一节　博弈表达的基本式

完全信息静态博弈是最简单的一种游戏[①]形式。在这种游戏里，每个参赛者同时出招一次，随即游戏结束。换言之，每一个参赛者各有一组可用的策略(行动)，他们选择某一行动后，游戏结束，而所有参赛者都得到各自相应的收益。这种形式的游戏可用三个基本的要素来完整地加以描述：

1. 参加博弈的有哪些“人”。人之所以加引号，是因为有些参与者可

① 为了轻松一些，不至于太严肃，有时会用游戏来替代“博弈”一词，在英语中博弈(Game)就是游戏。

能是企业、利益集团(例如卡特尔)、国家或是军事同盟。通常我们把这些参与者定义为阿拉伯数字 1,2,3,…,n,表示第一位参赛者、第二位参赛者等等,一直到第 n 位参赛者。所有参与者的集合定义为大写 I,即 I={1,2,…,n}。

2. 每一个参与者的可用策略。将每一个参赛者所有可用的策略集合定义为策略空间,用大写的 S_i 来表示。例如,第 i 位参赛者所有可用的策略为 x,y,z,那么 $S_i=\{x,y,z\}$,具体的策略定义为 s_{ij},它表示第 i 位参与者的第 j 个策略。显然 $s_{i1}=x, s_{i2}=y, s_{i3}=z$。策略空间是参与者所有可以选择的策略,不在策略空间之内的策略是参与者无法选择的策略。

3. 收益函数。每一个参与者得到多少收益取决于所有参与者的策略组合。策略组合与收益之间的关系,用收益函数 $u(s_1,s_2,\cdots,s_n)$来表示。收益函数表示的是在什么样的策略组合下每位参与者得到什么样的收益。所谓策略组合,就是所有参与者所选策略组成的一个 n 维向量。n 为参与者的个数。例如,一个博弈有三位参与者,某个策略组合为 $s=(s_{1j},s_{2k},s_{3h})$,即第 1 位参与者所用的策略为他的策略空间中的第 j 个策略,第 2 位参与者所用的策略为他的策略空间中的第 k 个策略,第 3 位参与者所用的策略为他的策略空间中的第 h 个策略。从数学来看,策略组合只不过是策略集合 $S_1\times S_2\times S_3$ 中的一个元素,即 $s\in S_1\times S_2\times S_3$。它表示第 1 位参与者只能在他的策略空间中选择策略,第 2 位参与者只能在他的策略空间中选择策略,第 3 位参与者也只能在他的策略空间中选择策略。如果博弈有 n 位参与者,那么策略组合 $s\in S_1\times S_2\times\cdots\times S_n$。在数学中,把 $S=S_1\times S_2\times\cdots\times S_n$ 称为 n 重笛卡尔积。

定义 2.1　博弈表达的基本式由参与者集合 I、策略空间 S 和收益函数 u 三个要素组成,即 $G=\{\mathrm{I},S,u\}$,其中 $\mathrm{I}=\{1,2,\cdots,n\}$,$S=\{S_1,S_2,\cdots,S_n\}$,$u=\{u_1,u_2,\cdots,u_n\}$。收益函数 $u_i:S\rightarrow R$,它表示第 i 位参与者在不同策略组合下所得到的收益。有时我们也把博弈表达的基本式写作 $G=\{S_1,S_2,\cdots,S_n;u_1,u_2,\cdots,u_n\}$。

下面我们用一些例子来说明如何将一个博弈转化为博弈的科学表达式——基本式。

例 2.1　石头、剪刀、布。

(1)参与者:孙悟空定义为1,猪八戒定义为2,I={1,2}。

(2)策略空间:孙悟空的策略空间为S_1={石头,剪刀,布},猪八戒的策略空间为S_2={石头,剪刀,布}。策略$s_{11}=s_{21}$=石头,$s_{12}=s_{22}$=剪刀,$s_{13}=s_{23}$=布。

(3)收益函数:$u_1(s_{1j},s_{2k})$和$u_2(s_{2j},s_{2k})$分别为孙悟空和猪八戒的收益函数。根据收益函数,我们可写出孙悟空和猪八戒在所有策略组合下可得的收益。例如,$u_1(s_{11},s_{21})=u_2(s_{11},s_{21})=0$,即孙悟空出石头,猪八戒也出石头,两者平局,收益都为0;$u_1(s_{11},s_{22})=+U$,$u_2(s_{11},s_{22})=-U$,即孙悟空出石头,猪八戒出剪刀,孙悟空休息,收益为+U,而猪八戒找水,收益为-U。当考察了所有的策略组合,并得到相应的收益,我们就可将其列成一个矩阵,如图2.1所示。由于博弈矩阵清楚地概括出了博弈的以上三个基本要素,因而博弈的基本式通常被认为等价于博弈矩阵。

		猪八戒 石头	剪刀	布
	石头	0,0	+U,-U	-U,+U
孙悟空	剪刀	-U,+U	0,0	+U,-U
	布	+U,-U	-U,+U	0,0

图2.1 石头、剪刀、布

不过,如果参与者超过三人,策略数增多,那么用博弈矩阵来表示博弈的基本式显然就不方便了。通常情况下,一个博弈的策略组合的数量等于k^n,k为每一个参与者拥有的策略数量,而n为参与者数量。[①]显然随着策略数增多或是参与者增多,博弈可能出现的结果呈几何数增长。从这一点来看,并不是所有的博弈基本式都可用博弈矩阵来概括。

例2.2 囚徒困境。

(1)参与者:囚徒1定义为1,囚徒2定义为2,I={1, 2}。

① 如果参与者拥有的策略数不相同,那么就是$k_1\times k_2\times\cdots\times k_n$,下标表示第几位参与者。

（2）策略空间：$S_1=$\{沉默，招供\}和 $S_2=$\{沉默，招供\}。策略 $s_{11}=s_{21}$ $=$沉默，$s_{12}=s_{22}=$招供。

（3）收益函数：$u_1(s_{1j}, s_{2k})$和 $u_2(s_{2j}, s_{2k})$分别为囚徒 1 和囚徒 2 的收益函数。所有的收益如下：

$u_1(s_{11}, s_{21})=-1$　$u_1(s_{11}, s_{22})=-9$　$u_1(s_{12}, s_{21})=0$　$u_1(s_{12}, s_{22})=-6$

$u_2(s_{11}, s_{21})=-1$　$u_2(s_{11}, s_{22})=0$　$u_2(s_{12}, s_{21})=-9$　$u_2(s_{12}, s_{22})=-6$

列成矩阵形式，得图 2.2。

		囚徒 2 沉默	囚徒 2 招供
囚徒 1	沉默	−1，−1	−9，0
	招供	0，−9	−6，−6

图 2.2　囚徒困境

例 2.3　斗鸡博弈。

在美国，一些飞车党党徒为了表示勇敢，通常由两个人分别驾驶两辆摩托车（或是汽车）从相反的方向急速对撞。如果有一方怕死而首先让道，那么让道的人就被称为小鸡(Chicken)。如果两人都同时让道，两人都被称为小鸡。被称为小鸡的人，在飞车党内地位低下，备受歧视。而不怕死的人，会在飞车党内享受崇高荣誉，受到众人推崇。显然，在斗鸡博弈中，如果两人都不让道，有可能都下地狱，或是不死也残废，收益为 $-\omega$；如果主动让道，又会受到大家的嘲笑，做人抬不起头，收益为 -10；如果不让道成为勇敢者，会受到大家尊敬，收益为$+10$。现在我们把斗鸡博弈转化为博弈的基本式。

（1）参与者：党徒 1 定义为 1，党徒 2 定义为 2，$I=\{1, 2\}$。

（2）策略空间：党徒 1 的策略空间 $S_1=$\{让，撞\}，党徒 2 的策略空间 $S_2=$\{让，撞\}。策略 $s_{11}=s_{21}=$让，$s_{12}=s_{22}=$撞。

（3）收益函数：$u_1(s_{1j}, s_{2j})$和 $u_2(s_{1j}, s_{2j})$为党徒 1 和党徒 2 的收益函数。所有的收益如下：

$u_1(s_{11}, s_{21})=-10$　$u_1(s_{11}, s_{22})=-10$　$u_1(s_{12}, s_{21})=+10$　$u_1(s_{12}, s_{22})=-\omega$

$u_2(s_{11},s_{21})=-10 \quad u_2(s_{11},s_{22})=+10 \quad u_2(s_{12},s_{21})=-10 \quad u_2(s_{12},s_{22})=-\omega$

列成矩阵形式，得图 2.3。

		党徒 2	
		让	撞
党徒 1	让	-10，-10	-10，+10
	撞	+10，-10	$-\omega$，$-\omega$

图 2.3 斗鸡博弈

例 2.4 古诺双头垄断模型。

在一个行业内只有两个生产者，这个行业的市场结构是一个寡头垄断的市场结构。垄断企业甲和垄断企业乙生产同质的产品，它们供应市场的产量分别为 q_1 和 q_2，市场中该产品的总供给为两个垄断企业的产量之和，即 $Q=q_1+q_2$，市场总需求为 $P(Q)=a-Q$（如果 $Q>a$，那么 $P(Q)=0$）。设垄断企业甲和垄断企业乙生产产量 q 的总成本为 $C(q)=cq$，即两个垄断企业不存在固定成本，且生产每单位产品的边际成本为常数 c，这里我们假定 $a>c$。

显然，两个垄断企业的策略（即产量）存在着相互依存性，彼此存在利益上的冲突，双方的利润都会受到对方所选产量的影响。垄断企业甲不能随心所欲地操纵市场，必须考虑到垄断企业乙可能作出的反应。对于垄断企业乙也同样如此。因而古诺模型是一个典型的双头博弈。现在我们将其转化为博弈的基本式。

（1）参与者：将垄断企业甲定义为 1，垄断企业乙定义为 2，$I=\{1, 2\}$。

（2）策略空间：垄断企业甲的策略空间为 $S_1=[0, a]$，垄断企业乙的策略空间为 $S_2=[0, a]$。策略 $s_1=q_1\in[0, a]$，$s_2=q_2\in[0, a]$。

（3）收益函数：垄断企业甲和乙的收益函数 $u(s_1, s_2)$ 实际上就是垄断企业的利润函数。收益即为利润。根据已知条件可知：

垄断企业甲的利润函数：$\pi_1(q_1,q_2)=q_1[a-(q_1+q_2)-c]$；

垄断企业乙的利润函数：$\pi_2(q_1,q_2)=q_2[a-(q_1+q_2)-c]$。

在古诺双头垄断模型中，由于策略、收益是连续的，因而不可能用

矩阵形式来表述该博弈，但这并不妨碍我们得到博弈的基本式。

通过上面不同的例子，我们已经看到，所有完全信息静态博弈都可以转化为基本式。有了对博弈的科学表述，随后的问题就是如何找到博弈的均衡解，而纳什均衡这个概念构成了博弈均衡解的基础。

第二节　纳什均衡

在介绍纳什均衡之前，我们先从简单的开始，然后逐步深入，最终给出纳什均衡的定义。由于纳什均衡的重要性，在本章的最后，我们还会给出纳什均衡的一个非严格的数学证明。

一、严格优策略均衡与一般均衡

研究博弈理论最重要的目的就是希望在面对一个博弈时，可以了解并预测博弈最后的结果。读者在学习经济学的时候，已经看到在均衡状态下，所有行为主体都没有积极性主动偏离均衡状态，即使偏离了均衡状态，各行为主体也会想方设法回到均衡状态，因为均衡对所有各方来说是最优的。这说明一个博弈最后的结果最有可能是一个均衡状态。在博弈中，最简单的均衡概念就是严格优策略均衡。

我们还是用囚徒困境来加以说明。在囚徒困境中，什么是博弈最可能出现的结果，或换句话说，博弈的均衡解是什么？通过图 2.2，我们可以看出，如果囚徒 2 选择沉默，那么囚徒 1 最优的策略是招供；如果囚徒 2 选择招供，那么囚徒 1 最优的策略还是招供。这说明无论囚徒 2 的策略是什么，囚徒 1 选择招供都比选择沉默要优。对于囚徒 2 而言，同样的分析也成立。因此，囚徒困境最可能出现的结果是(招供，招供)。没有哪个囚徒会主动偏离这个结果，因为偏离就意味着要遭受更大的损失。例如，囚徒 2 偏离它，选择沉默，那么坐牢就不是 6 个月，而是 9 个月。因此，(招供，招供)也是囚徒困境的均衡解。

定义 2.2 在博弈 G={I, S, u}中，在其他参与者任意①给定的策略组合下，即∀ $s_{-i}=(s_1,\cdots,s_{i-1},s_{i+1},\cdots,s_n)\in S_{-i}=(S_1,\cdots,S_{i-1},S_{i+1},\cdots,S_n)$，参与者 i 存在一个策略 s_i^* 使得对于∀ $s_i\in S_i$，$s_i^*\neq s_i$，都有：

$$u_i(s_1,\cdots,s_i^*,\cdots,s_n)>u_i(s_1,\cdots,s_i,\cdots,s_n)$$

我们就称 s_i^* 为参与者 i 的严格优策略。

在囚徒困境中，无论囚徒 2 的策略是什么，囚徒 1 选择招供所得到收益都要大于选择沉默，所以招供是囚徒 1 的严格优策略。对于囚徒 2 而言，无论囚徒 1 选择什么策略，囚徒 2 选择招供都要优于沉默，所以囚徒 2 的严格优策略为招供。显然，(招供，招供)构成了囚徒困境的一个均衡解。通过对囚徒困境的分析，很容易得到一个一般性的结论，即如果一个博弈中的所有参与者都存在严格优策略，那么严格优策略组合一定是该博弈的惟一均衡解。

命题 2.1 如果在一个博弈基本式 G={I, S, u}中，每个参与者都存在严格优策略，那么严格优策略组合一定是该博弈惟一的均衡解。

证：假设一个博弈为 G={I, S, u}，其中每一个参与者都存在严格优策略。根据定义可知严格优策略具有惟一性，如果不是那就意味着某个参与者 i 至少有两个严格优策略，不妨设它们分别为 s_i^* 和 $s_i\#$，根据严格优策略的定义，可知：

$$u_i(s_i^*,\ s_{-i})>u_i(s_i\#,\ s_{-i})$$
$$u_i(s_i\#,\ s_{-i})>u_i(s_i^*,\ s_{-i})$$

显然这是自相矛盾的。因而严格优策略具有惟一性。由于严格优策略具有惟一性，因而每一个参与者假如不选择严格优策略，将意味着他得到的收益不为最优，显然这与参与者是理性的相矛盾，从而命题得证。

定义 2.3 在博弈 G={I, S, u}中，如果对所有参与者 i 而言，都存在 s_i^* 为严格优策略(其中∀ $i\in I$)，那么策略组合 $s^*=(s_1^*,\cdots,s_i^*,$

① 在数学中，“任意”用∀ 来表示，它为 Any(任意)头一个字母的倒写，“存在一个”用∃来表示，它为 Exist(存在)头一个字母的倒写。

$\cdots,s_n^*$)就称为严格优策略均衡。

在有些情况下，优策略也可能形成均衡，而不一定必须是严格优策略，但也可能出现多重均衡。

定义 2.4　在博弈 $G=\{I,S,u\}$ 中，如果对参与者 i 而言，在其他参与者任意给定的策略组合下，即 $\forall\ s_{-i}=(s_1,\cdots,s_{i-1},s_{i+1},\cdots,s_n)\in S_{-i}=(S_1,\cdots,S_{i-1},S_{i+1},\cdots,S_n)$，存在一个策略 s_i^* 使得对于 $\forall\ s_i\in S_i$，都有：

$$u_i(s_1,\cdots,s_i^*,\cdots,s_n)\geqslant u_i(s_1,\cdots,s_i,\cdots,s_n)$$

并且至少存在一个 s_i 使得上述不等式严格成立。我们就称 s_i^* 为参与者 i 的优策略。

如果参与者只存在优策略而不是严格优策略，那么一个可能的后果是博弈不存在惟一的均衡解，从而使得我们无法对博弈的结果作出明确的预测。极端的情况是博弈有可能出现各种各样的结果，这时理论的力量就丧失掉了。

在读者学习微观经济学的时候，已知证明一般均衡存在的一个不可缺少的重要前提是市场结构为完全竞争的市场结构。完全竞争的本质是所有的市场参与者—— 消费者和生产者都是价格的接受者，而不是价格的影响者。无论是消费者还是生产者，他的决策(商品的消费量和供给量)对市场价格的影响都趋于零，从而每一个参与者的策略都不会影响其他参与者的策略，也就是参与者之间不存在策略依存性，而这种特征和严格优策略是完全相同的。对消费者而言，他是在给定商品价格、收入等的条件下使自身的效用最大化，至于其他参与者如何行动对他的决策不产生影响；对生产者而言，他是在给定商品价格、要素价格、成本、技术等的条件下使自身的利润最大化，至于其他参与者如何，同样与他不存在联系。因而完全竞争市场上的每一个参与者实际上作出的策略都是严格优策略，形成的一般均衡就是一个严格优策略均衡。如果有的参与者作出的策略是优策略，而不是严格优策略，例如在完全竞争下厂商的利润为零时就会出现这种情况，那么一般均衡就可能是多重的。

一般均衡为严格优策略均衡，大大加深了我们对一般均衡和博弈论的理解，由于一般均衡只不过是博弈论中纳什均衡的一个特例，因而

博弈论适用的领域就要比一般均衡所描绘的范围广，这就是为什么博弈论能够在所谓"市场失灵"——外在性、公共产品、垄断、道德风险、逆向选择的世界大展拳脚，制度、产权这些传统经济学作为外生给定因素来处理的东西，反而成为博弈论研究中的考察对象。

在所有博弈中存在严格优策略的博弈毕竟是少数，绝大多数博弈并不存在严格优策略。不过在有些博弈中，虽然没有严格优策略但却存在严格劣策略，通过剔除严格劣策略仍可能得到博弈的均衡解。

定义 2.5 在博弈 $G=\{I, S, u\}$ 中，s_i' 和 s_i'' 为参与者 i 的两个可行策略，即 $\exists\ s_i', s_i'' \in S_i$。如果对于 $\forall\ s_{-i} \in S_{-i}$，有 $u_i(s_i', s_{-i}) > u_i(s_i'', s_{-i})$，那么 s_i'' 就称相对于策略 s_i' 的严格劣策略。

作为理性者，显然没有哪个参与者会选择严格劣策略，因而我们可以将严格劣策略从参与者的策略空间中剔除。通过不断地剔除严格劣策略就可能最终得到博弈的均衡解。下面我们用例子来加以说明。

考虑图 2.4 所表示的抽象博弈。

		参与者 2 左	参与者 2 中	参与者 2 右
参与者 1	上	1,0	1,2	0,1
	下	0,3	0,1	2,0

图 2.4 上下——左中右

在这个博弈中，参与者 1 显然不存在严格优策略，也不存在严格劣策略。参与者 2 虽不存在严格优策略，但却存在严格劣策略。策略右相对于策略中就是一个严格劣策略。因为无论参与者 1 选择什么策略，参与者 2 选择中都要优于选择右。如果参与者是理性的，无论什么情况，参与者 2 绝对不会选择右，因而我们可以将右从参与者 2 的策略中剔除，从而得到图 2.5。

		参与者 2 左	参与者 2 中
参与者 1	上	1,0	1,2
	下	0,3	0,1

图 2.5 上下——左中

从图 2.5 可以看到，对于参与者 1 而言，下是一个严格劣策略，可从参与者 1 的策略空间中剔除，最终得到图 2.6。

		参与者 2	
		左	中
参与者 1	上	1,0	1,2

图 2.6　上——左中

显然(上，中)就是该博弈惟一的均衡解。

这种方法在博弈论中被称为重复剔除严格劣策略。

存在严格劣策略，自然也存在劣策略，它的定义类似于优策略的定义，请读者自己举一反三，这里不再说明。重复剔除严格劣策略这种方法能不能运用到劣策略呢？答案是否定的。因为重复剔除劣策略有可能将博弈的均衡解给剔除掉，只有在严格劣策略下，才不会出现这种情况。原因在于，严格劣策略无论在什么情况下它都使严格不等式成立，而在劣策略下有些策略的收益是相等的。

严格劣策略虽然要求的条件没有严格优策略那么苛刻(严格优策略要求对所有策略都要优，而严格劣策略只是相对于某个策略而言，这意味着存在严格优策略必然存在严格劣策略，存在严格劣策略却不一定存在严格优策略)，但毕竟具有严格劣策略的博弈仍然是少数，对于更为一般的博弈，如何求出均衡解呢？这就引出了“纳什均衡”这个概念。

二、纳什均衡

在前面我们举了许多例子，例如，诺曼底登陆、鸽派和鹰派、富人穷人博弈、斗鸡博弈。在这些例子中，没有哪一个参与者拥有严格优策略或是严格劣策略，但从直觉上来说，我们都认为这些博弈存在着均衡解。例如，富人穷人博弈的均衡解是(巡逻，不巡逻)，鸽派和鹰派的均衡解是两个(鸽派政策，鹰派政策)和(鹰派政策，鸽派政策)。我们再举一例。

例 2.5　上中下——左中右。

这个博弈的基本式如图 2.7 的博弈矩阵所示。

		参与者 2		
		左	中	右
参与者 1	上	5,3	0,4	3,5
	中	4,0	5,5	4,0
	下	3.5	0,4	5,3

图 2.7 上中下——左中右

显然这个博弈不存在着严格劣策略，但它存在均衡解。

定义 2.6 一个策略组合 $s^*=(s_1^*,\cdots,s_n^*)$ 是博弈 $G=\{I,S,u\}$ 的一个纳什均衡，如果对于每一个参与者 $i=1,2,\cdots,n$，对于所有的 $s_i\in S_i$，不等式

$$u_i(s_i^*,s_{-i}^*)\geqslant u_i(s_i,s_{-i}^*)$$

都成立。

理解纳什均衡的关键是抓住两点：一是每一个参与者，二是所有的策略。

上述定义的思想十分简单，如果一个策略组合能够同时使得所有人都得到（相对其他人的策略）最好的结果，那么这个策略组合一定是纳什均衡。利用这个思想我们就可以对各种各样的博弈进行分析，并找出它们的均衡解。

在例 2.5 中，纳什均衡为（中，中），为什么？如果给定参与者 1 的策略为中，那么参与者 2 最优的策略为中；反过来，给定参与者 2 的策略为中，那么参与者 1 最优的策略为中，所以（中，中）这个策略组合为纳什均衡。其他的策略组合都不是纳什均衡。例如，策略组合（上，右）不是纳什均衡，因为给定参与者 1 的策略为上，参与者 2 最优的策略为右，但给定参与者 2 的策略为右，参与者 1 的最优策略却为下。

根据纳什均衡这个概念，我们来看看前面提到的一些例子的均衡。

例，鸽派和鹰派（见图 1.3）。

如果美国的策略是鸽派政策，那么苏联的最优策略是鹰派政策；反过来，如果苏联的策略是鹰派政策，那么美国的最优策略是鸽派政策，

所以(鹰派政策,鸽派政策)是该博弈的一个纳什均衡。如果美国的策略是鹰派政策,那么苏联的最优策略是鸽派政策;如果苏联的策略为鸽派政策,那么美国的最优策略为鹰派政策,所以(鸽派政策,鹰派政策)也是该博弈的一个纳什均衡。

例,富人穷人博弈(见图 1.5)。

如果给定穷人的策略是不巡逻,那么富人的最优策略是巡逻;反过来,如果给定富人的策略是巡逻,那么穷人的策略是不巡逻,所以(巡逻,不巡逻)是富人穷人博弈的纳什均衡。

例,斗鸡博弈(见图 2.3)。

如果党徒 2 不要命地撞过来,那么党徒 1 最优的策略是让,失去尊严总比命丧九泉要好;如果党徒 1 贪生怕死,那么党徒 2 最优的策略就是撞(因为党徒 1 会让),所以(让,撞)就是斗鸡博弈的一个纳什均衡。另一个纳什均衡为(撞,让)。

现在我们来看囚徒困境。囚徒困境是一个包含严格优策略均衡的博弈。在囚徒困境中,(招供,招供)是一个严格优策略均衡,但它也是一个纳什均衡。因为给定囚徒 2 的策略为招供,那么囚徒 1 最优的策略为招供;反过来,如果给定囚徒 1 的策略为招供,那么囚徒 2 的最优策略为招供,所以(招供,招供)为纳什均衡。这表明严格优策略均衡必定是纳什均衡,反之不成立。例如,鸽派和鹰派博弈、富人和穷人博弈、斗鸡博弈、上中下——左中右博弈就不存在严格优策略,但存在纳什均衡。

命题 2.2 博弈 $G=\{I,S,u\}$ 存在严格优策略均衡 $s^*=(s_1^*,\cdots,s_n^*)$,那么它一定是一个纳什均衡,但反之不成立。

证:利用反证法。

已知策略组合 $s^*=(s_1^*,\cdots,s_n^*)$ 是博弈 $G=\{I,S,u\}$ 的严格优策略,但不是纳什均衡。我们不妨假设纳什均衡为 $s\#=(s_1\#,\cdots,s_n\#)$。根据严格优策略的定义可知,一定存在参与者 $i=1,2,\cdots,n$,有:

$$u_i(s_i^*,s_{-i}\#)>u_i(s_i\#,s_{-i}\#)$$

而根据纳什均衡的定义,一定有:

$$u_i(s_i\#,s_{-i}\#)\geq u_i(s_i^*,s_{-i}\#)$$

显然这是自相矛盾的。

要证第二个观点，我们不妨假设凡是纳什均衡一定是严格优策略均衡。只要找出一个反例，这个假设就不攻自破。已知前面的斗鸡博弈存在着纳什均衡，但不存在严格优策略均衡，这意味着纳什均衡一定是严格优策略均衡是错误的。命题得证。

再看例子上下——左中右博弈(见图 2.4)。这个博弈存在严格劣策略，但不存在严格优策略。经过重复剔除严格劣策略后，我们得到的均衡解为(上，中)。显然这个均衡解也是纳什均衡。因为给定参与者 2 的策略为中，那么参与者 1 的最优策略为上；给定参与者 1 的策略为上，那么参与者 2 的最优策略为中，所以(上，中)也是纳什均衡。

命题 2.3 如果策略组合 $s^*=(s_1^*,\cdots,s_n^*)$为博弈 $G=\{I,S,u\}$的纳什均衡，它一定不会被重复剔除严格劣策略的方法剔除掉；如果进行了重复剔除严格劣策略后剩下惟一策略组合，那么这个策略组合一定是纳什均衡。

证：利用反证法。

如果策略组合 $s^*=(s_1^*,\cdots,s_n^*)$为博弈 $G=\{I,S,u\}$的纳什均衡，但却在重复剔除严格劣策略的过程中被剔除。这意味着在重复剔除严格劣策略的某个阶段，存在某个参与者 $i\in I$，有：

$$u_i(s_i\#,s_{-i})>u_i(s_i^*,s_{-i})，也就是$$

$$u_i(s_i\#,s_{-i}^*)>u_i(s_i^*,s_{-i}^*)成立。$$

显然这与 $s^*=(s_1^*,\cdots,s_n^*)$为博弈 $G=\{I,S,u\}$的纳什均衡相矛盾。

如果重复剔除严格劣策略后剩下惟一的策略组合不是纳什均衡，那就意味着纳什均衡在重复剔除严格劣策略的某个阶段已被剔除，显然这与我们前面证明的纳什均衡能够在重复剔除严格劣策略中保存下来相矛盾。命题得证。

命题 2.2 和命题 2.3 表明，纳什均衡包含着严格优策略均衡和重复剔除严格劣策略均衡，它是一个更为宽广、更为适用的均衡概念。

第三节　运用举例

现在我们就运用前面所讲的知识来具体分析一些经典的博弈模型。通过这些经典例子使读者进一步明确：(1)如何将一个博弈转化为基本式；(2)如何利用纳什均衡的概念求出博弈的均衡解。

一、古诺模型

最早对策略依存的经济现象进行研究，并提出均衡解的经济学模型就是古诺模型了。1838 年古诺提出了一个双头垄断模型。古诺的研究理所当然地成为博弈论研究中的经典文献之一，并开创了产业组织理论研究的先河。后人在古诺模型的基础上又发展出许多的变型。古诺模型的主要内容如下：

一个行业有两个寡头企业垄断，两个寡头企业生产的产品完全同质。令 q_1,q_2 分别表示企业 1 和企业 2 生产的同质产品的产量，市场中该产品的总供给为 $Q=q_1+q_2$，令 $P(Q)=a-Q$ 为需求曲线(假设如果 $Q>a$，那么 $P(Q)=0$)。设企业 $i(i=1,2)$生产 q_i 的总成本 $C_i(q_i)=cq_i$，即企业不存在固定成本，且生产每单位产品的边际成本为常数 c，这里我们假定 $c<a$。

为求出古诺博弈中的纳什均衡，我们首先要将其化为标准式的博弈。前面已讲过，博弈的标准式包含以下要素：(1)博弈的参与者；(2)每一个参与者的策略集合；(3)在每一种可能的策略组合下，每一个参与者的收益，即收益函数。在古诺模型中，参与者集合 $I=\{1,2\}$，而策略 s_i 就是每一个垄断企业的产量 q_i，策略集合 $S_i=[0,a]$。最后是垄断企业的收益函数，显然企业的利润函数就是垄断企业的收益函数。对于第 1 个企业，其收益函数 $u_1(s_1,s_2)$实际上为：

$$\pi_1(q_1,q_2)=q_1[p(q_1+q_2)-c]$$

对第 2 个企业来说，其收益函数 $u_2(s_1,s_2)$ 为：

$$\pi_2(q_1,q_2)=q_2[p(q_1+q_2)-c]$$

根据纳什均衡的定义，对于第 1 个企业[①] 必须满足：

$$\pi_1(q_1^*,q_2^*)\geqslant\pi_2(q_1,q_2^*)$$

它对所有的 $q_1\in[0,a]$ 成立。

这实际上等价于在给定第 2 个企业策略（产量）的情况下求第 1 个企业的利润最大化，即：

$$\max_{q_1\in S_1}\pi_1(q_1,q_2^*)$$

$$\max_{0\leqslant q_1\leqslant a}\pi_1(q_1,q_2^*)=\max_{0\leqslant q_1\leqslant a}q_1[a-(q_1+q_1^*)-c]$$

求解一阶条件，得：

$$q_1=\frac{1}{2}(a-q_2^*-c) \tag{2.1}$$

利用对称性，得：

$$q_2=\frac{1}{2}(a-q_1^*-c) \tag{2.2}$$

解这对方程组，得：

$$q_1^*=q_2^*=\frac{a-c}{3} \tag{2.3}$$

即，如果给定第 2 个企业的产量为 $(a-c)/3$，那么第 1 个企业的最优产量为 $(a-c)/3$；如果给定第 1 个企业的产量为 $(a-c)/3$，那么第 2 个企业的最优策略为 $(a-c)/3$，所以 $((a-c)/3,(a-c)/3)$ 为古诺模型的纳什均衡。

公式(2.1)和(2.2)又可以看作是垄断企业 1 和垄断企业 2 的策略函数，即企业 i 根据企业 j 的策略而定的行动计划。例如给定企业 2 的

① 对第 1 个企业的讨论完全适合第 2 个企业。

策略为 q_2，企业 1 的最优策略为 $q_1=\frac{1}{2}(a-q_2-c)$。公式(2.1)和(2.2)构成了一个方程组，我们定义这个方程组为 $R(q_1,q_2)$，即：

$$R(q_1,q_2)=\begin{cases}R_1(q_2)=\frac{1}{2}(a-c-q_2)=q_1\\ R_2(q_1)=\frac{1}{2}(a-c-q_1)=q_2\end{cases}$$

显然这个方程组是 $R:S_1\times S_2\to S_1\times S_2$ 的一个映射。它的定义域为策略组合，值域也为策略组合。通过前面的分析我们已经知道存在纳什均衡 (q_1^*,q_2^*)，它使得 $R(q_1^*,q_2^*)=\begin{cases}q_1^*\\ q_2^*\end{cases}$，即方程组将策略组合 (q_1^*,q_2^*) 映射成它自身，在数学中把映射成它自身的点称为不动点。[①] 例如 $f(x)=x$，x 这个点就被称为不动点。两个寡头企业的策略函数也可用图 2.8 表示。

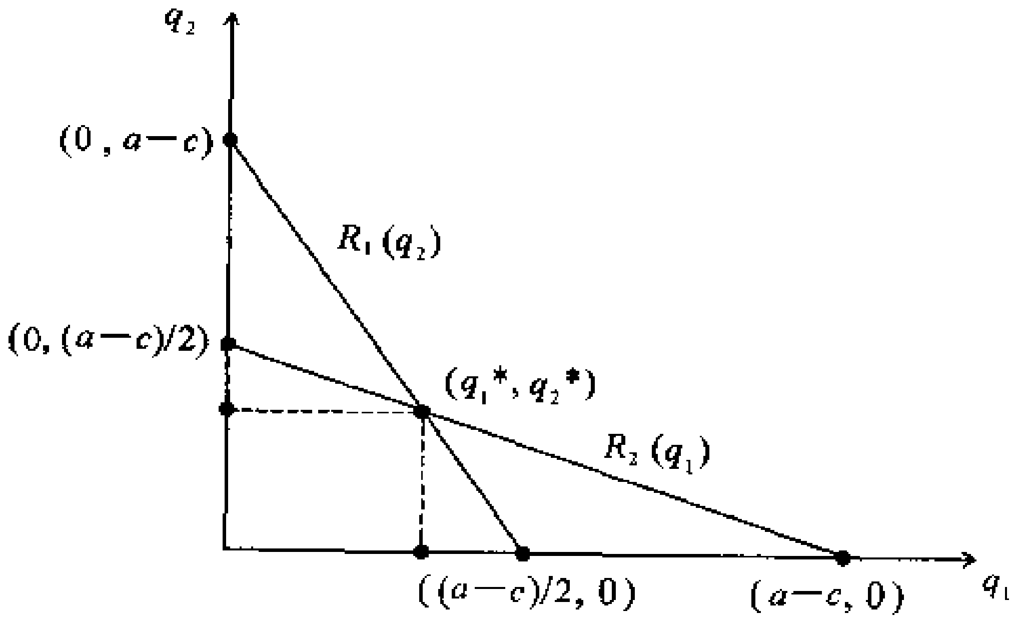


图 2.8　古诺模型与不动点

$R_1(q_2)$ 和 $R_2(q_1)$ 分别为垄断企业 1 和 2 的策略函数，两者的交点

① 数学中的点就是向量。

即为纳什均衡，也是不动点。

古诺模型的意义在于指明垄断企业之间的串谋是不稳定的，每一个企业都有欺骗对方的动机。要说明这一点，我们假设两个垄断企业串谋在一起成立所谓的卡特尔(相当于一个企业)。根据前面的假设，这个卡特尔的最大利润是：

$$\max_{0\leq Q\leq a} \pi(Q)=pQ-cQ$$
$$\pi(Q)=Q(a-Q-c)$$

求一阶条件，得：

$$Q=\frac{a-c}{2}$$

这表明如果两个垄断企业联合起来，限制总产量为$(a-c)/2$(非卡特尔下总产量为$2(a-c)/3$)，就能使两个企业总利润最大。由于两个垄断企业的成本完全一样，各自生产$\frac{a-c}{4}$的产量就能均分利润。每个企业所得利润为$(a-c)^2/8$，大于非卡特尔下的利润$(a-c)^2/9$。但问题是，如果给定$q_2=\frac{a-c}{4}$(即第2个企业遵守卡特尔协议)，那么根据(2.1)式第1个企业的最优产量为$q_1=\frac{3}{8}(a-c)$，显然大于卡特尔所制定的产量$\frac{a-c}{4}$。对于第2个企业同样也存在欺骗的动机，因而卡特尔是不稳定的，串谋双方都在试图欺骗对方。实际上，古诺模型所揭示出的卡特尔困境只不过是囚徒困境的另一种表现。

二、两党政治

西方发达资本主义国家在政治制度上基本都实行多党制。从政治主张来看，可以划分为保守主张和激进主张。由于党派之间的斗争存在明显的策略依存性，因而多党制竞争是一种博弈过程，我们可以用所学的博弈论知识来抽象概括这种多党制竞争。为了简单，我们这里只考虑两党制。在章后习题中会要求读者求解三党制的纳什均衡。

假设保守主张用 0 来表示，激进主张用 1 来表示。选民在保守主张和激进主张之间均匀分布，即在 0 和 1 之间均匀分布。两党需要确定某种政治主张以吸引选民投它的票，获得的选票越多越好。我们用图 2.9 来说明两个党派如何争夺选民。

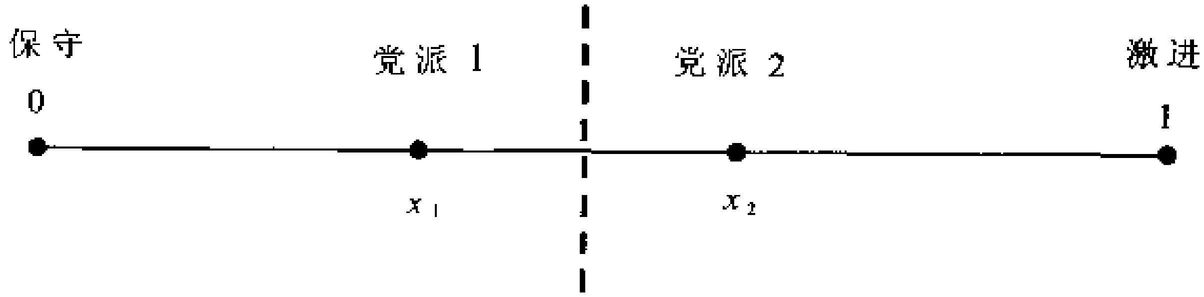


图 2.9　两党争夺选民

如果党派 1 的政治主张(策略)为 x_1，它较偏向于保守主张，那么它能获得保守选民的选票，即得到分布在[0,x_1]的选民的支持。由于我们假设选民均匀分布在[0,1]的区间，所以它获得的投票数可用 x_1 来表示。对于党派 2 而言，它的政治主张为 x_2，较偏向激进主张，所以它能获得激进选民的选票，即获得分布在[x_2,1]的选民的支持。他所获得的投票数可用 $1-x_2$ 来表示。由于选民平均分布在[0,1]之间，所以分布在[x_1,x_2]的选民的选票被两个党派平分。图 2.9 中的虚线的左边为支持党派 1 的选民，虚线的右边为支持党派 2 的选民。显然，在给定党派 1 的策略 x_1，党派 2 的策略 x_2 越接近 x_1，获得的选民就会越多；对于党派 1 同样存在这样的情况。那么在这个两党政治中纳什均衡是什么？

(1)参与者：党派 1 定义为 1，党派 2 定义为 2，I={1,2}。

(2)策略空间：党派 1 的策略空间为 $S_1=[0,1]$，策略 $x_1\in[0,1]$；党派 2 的策略空间为 $S_2=[0,1]$，策略为 $x_2\in[0,1]$。

(3)收益函数：两个党派的收益实际上就是两个党派所能获得的选民数，因而收益函数为两个党派的政治主张的函数。根据已知条件，党派 1 的收益函数 $u_1(x_1,x_2)$为：

$$u_1(x_1,x_2)=\begin{cases}\dfrac{x_1+x_2}{2} & x_1<x_2\\ \dfrac{1}{2} & x_1=x_2\\ 1-\dfrac{x_1+x_2}{2} & x_1>x_2\end{cases}$$

党派 2 的收益函数 $u_2(x_1,x_2)$ 为：

$$u_2(x_1,x_2)=\begin{cases}\dfrac{x_2+x_1}{2} & x_2<x_1\\ \dfrac{1}{2} & x_2=x_1\\ 1-\dfrac{x_2+x_1}{2} & x_2>x_1\end{cases}$$

上述收益函数的思想相当简单：如果党派 1 偏向保守主张，即 $x_1<x_2$，那么它得到 $(x_1+x_2)/2$ 的选票；如果它和党派 2 的政治主张一样，那么双方各得 1/2 的选票；如果党派 1 偏向激进主张，即 $x_1>x_2$，那么它得到 $1-(x_1+x_2)/2$ 的选票。对于党派 2，其收益函数同理。

两个党派显然都力求使它的收益最大化，即

$$\max_{x_i\in[0,1]} u_i(x_i,x_j^*) \qquad 其中\ i,j=1,2,i\neq j\ 。$$

当 $x_1\neq x_2$ 时，党派 1 的最优策略是 $x_1\to x_2$，即党派 1 的政治主张越向党派 2 靠拢，"争取"到的选票越多；而党派 2 的最优策略也是 $x_2\to x_1$，即党派 2 的政治主张越向党派 1 靠拢，"争取"到的选票越多。因而，无论党派 1 还是党派 2，他们收益最大化的必要条件是 $x_1=x_2$。在 $x_1=x_2$ 时，党派 1 和党派 2 分别获得 1/2 的选票。但是，当 $x_1=x_2\neq 1/2$，即不位于中点（代表着政治上的中间路线）时，党派 1 和党派 2 都有积极性向中点移动（意味着政治主张趋于中间路线）。因为如果党派 j 的策略不变，即 $x_j\neq 1/2$，党派 i 向中点移动（奉行中间路线）至少可获得 $1/2+|1/2-x_i|/2$ 的选民，显然高于 1/2。所以只有当 $x_1=x_2$ 并位于中点时，党派 1 和党派 2 才没有动机偏离。因此，中点是两党政治惟一

的纳什均衡。

两党政治这个模型虽然简单，但却深刻地说明了西方发达国家无论什么党派上台，长期奉行的政治路线通常为中间路线，即不过分保守，也不过分激进。像美国、英国这些两党制的国家就是典型。此外，它还揭示出各党派的政治主张具有趋同性，奉行中间路线的党派最有可能获胜，这意味着真正的政治改革通常难以实行，日本就是一个很好的例子。

两党政治这个模型还可以进一步变型，例如三党制。在三党制下，纳什均衡非惟一，而是有无穷多的可能均衡，但最为核心的东西并没有变，那就是没有哪一个党派能够获得绝对多数选民的支持，因而通常的政府和议会是以党派联合执政的方式出现的，这就解释了为什么实行多党制（$i>2$）的国家通常是联合党执政，而很少出现像美国那样的情况。

三、公共地悲剧

18 世纪初期，资本主义生产方式在英国得到了确立，并开始迅速发展。而它的最初动力是纺织品市场需求的日益扩大，使得毛纺工业迅速发展，纺织机、蒸汽机相继得到运用，直接导致第一次工业革命。羊毛作为原料，市场需求巨大。但当时英国农村的土地属于公社所有，牧场更是如此。在这种情况下，土地公有实际上成为有效生产羊毛的一个重要障碍，最终导致马克思称之为“羊吃人”的圈地运动。1739 年，休谟（Hume）在其所写的一篇文章中指出，公共牧场实际上会导致过度放牧，从而为英国圈地运动的产生提供了一种说明。在经济学中，公共地悲剧现在已成为说明私有制存在理由的经典模型之一。

公共地悲剧讲述的是一个村社有 $n\geqslant 2$ 个牧民。村社的公共牧场向每一个村社的成员开放。在春天，牧民同时选择所需要饲养的羊群数。假设羊群数是连续可分的。每年夏季，所有的牧民都在村社的公共牧地上放牧自己的羊群。如果第 i 个牧民拥有的羊群数为 g_i，那么村社所有的羊群数就为 $G=g_1+\cdots+g_n$。买和饲养一只羊的成本为 c，这里我们假设它是一个常数。每只羊的价值为羊的总头数的函数，即 $v(G)$。由于

每只羊都需要一定的绿地才能生存，因而村社的牧场所能供养的羊群数就有一个上限，我们定义为 G_{max}，并假设当所饲养的羊群数超过牧场所能供养的上限时，每只羊的价值为零（即羊群过多造成牧场退化，由于没有足够的饲料，从而引起羊群死亡）。用公式表示为：$v(G)>0$，$G<G_{max}$；$v(G)=0$，$G\geqslant G_{max}$。由于一开始饲养的羊群数很少，因而增加一只羊的数量对已经饲养的羊来说不会造成太大的伤害，但这种伤害是递增的，即当 $G<G_{max}$ 时，$v'(G)<0$，$v''(G)<0$。一阶导数为负表明，随着羊的数量增加，羊的价值会下降；二阶导数为负则表明，随着羊的数量增加，羊的价值下降是递增的。这种关系可用图 2.10 表示。

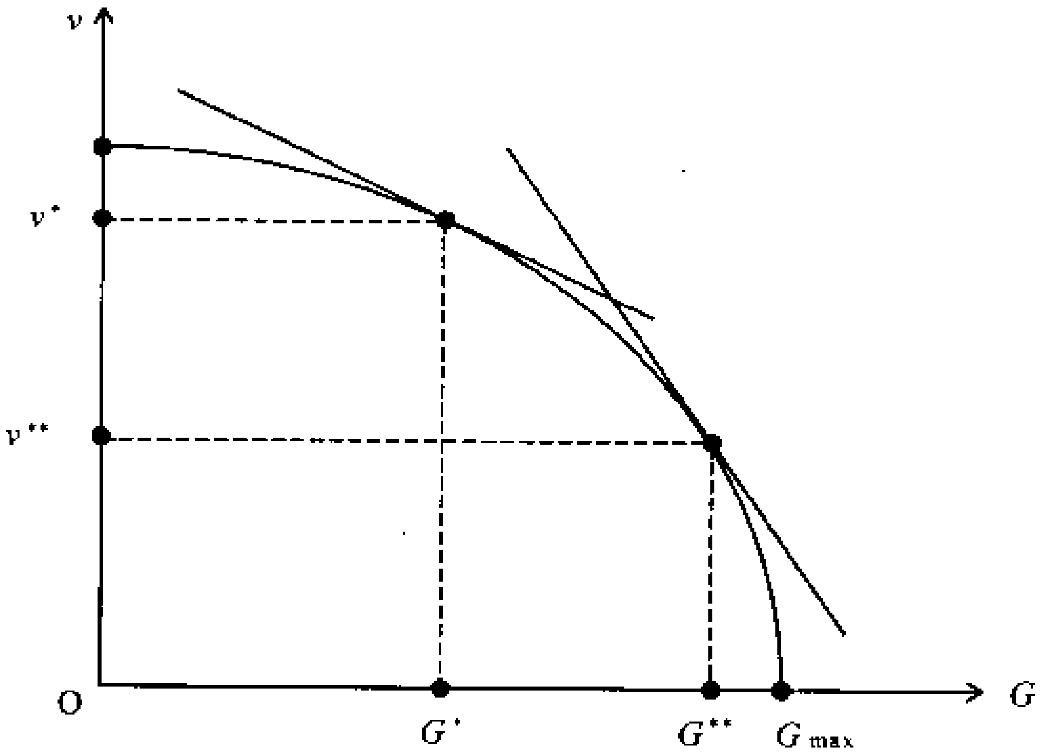


图 2.10　羊的价值以递增的速度下降

现在将该博弈转化为基本式。

（1）参与者：定义牧民为 $1,2,\cdots,n$，$I=\{1,2,\cdots,n\}$。

（2）策略空间：牧民 i 的策略就是选择 g_i，策略空间 $S_i=[0,G_{max})$。

（3）收益函数：当其他的牧民所放牧的羊群数为 $(g_1,\cdots,g_{i-1},g_{i+1},\cdots,g_n)$ 时，牧民 i 从放牧 g_i 的羊群中所获得的收益为：

$$u_i(g_1,\cdots,g_i,\cdots,g_n)=g_iv(g_1+\cdots+g_{i-1}+g_i+g_{i+1}+\cdots+g_n)-cg_i \quad (2.4)$$

每个牧民追求收益最大化，即：

$$\max_{g_i} u_i(g_1,\cdots,g_n)$$

假如$(g_1^*,\cdots,g_n^*)$是纳什均衡的话，对于每一个牧民i而言，在给定其他人的策略组合$(g_1^*,\cdots,g_{i-1}^*,g_{i+1}^*,\cdots,g_n^*)$下，$g_i^*$一定会使(2.4)式最大化。

最大化的必要条件(一阶条件)为：

$$v(g_i+g_{-i}^*)+g_iv'(g_i+g_{-i}^*)-c=0 \tag{2.5}$$

最大化的充分条件(二阶条件)为：

$$2v'(g_i+g_{-i}^*)+g_iv''(g_i+g_{-i}^*)<0$$

其中$g_{-i}^*=g_1^*+\cdots+g_{i-1}^*+g_{i+1}^*+\cdots+g_n^*$。利用对称性，可知(2.5)式是每一个牧民的收益最大化的充要条件。根据纳什均衡的定义，g_i^*一定是(2.5)式的解。将g_i^*代入(2.5)式，并将所有牧民的必要条件相加，得：

$$nv(G^*)+G^*v'(G^*)-nc=0 \tag{2.6}$$

用n除(2.6)式，得：

$$v(G^*)+\frac{1}{n}G^*v'(G^*)-c=0 \tag{2.7}$$

其中$G^*=g_1^*+\cdots+g_n^*$。(2.7)式是纳什均衡下，所有牧民所养羊之和必须满足的条件。

与私人利益最大化相对立的是社会的收益最大化，社会的目标是使下式最大化，即：

$$\max_{0\leqslant G<\infty} Gv(G)-Gc$$

一阶条件为：

$$v(G^{**})+G^{**}v'(G^{**})-c=0 \tag{2.8}$$

二阶条件为：

$$2v'(G^{**})+G^{**}v''(G^{**})<0$$

G^{**}就是社会的最优解。现在我们需要证明的是$G^{*}>G^{**}$，即公共地存在过度利用，怀海特（Whitehead）把这种情况称为“公共地悲剧”。

用反证法证明。假设$G^{*}\leqslant G^{**}$。

因为 $v'<0$

所以 $v(G^{*})\geqslant v(G^{**})$ （第 1 项）

又因为 $v''<0$

所以 $0>v'(G^{*})\geqslant v'(G^{**})$

所以 $|v'(G^{*})|\leqslant|v'(G^{**})|$ （第 2 项）

根据假设 $G^{*}\leqslant G^{**}$

所以 $\frac{G^{*}}{n}<G^{**}$ （第 3 项）

根据（第 1 项）、（第 2 项）、（第 3 项）比较（2.7）式和（2.8）式，可得出（2.7）式大于（2.8）式的结论。显然这与（2.7）和（2.8）都等于零相矛盾，所以$G^{*}>G^{**}$。命题得证。

$G^{*}>G^{**}$表明：第一，公共牧场被过度利用，资源没有达到有效配置；第二，拥有羊群越多的牧民损失越多，因为每只羊的价值$v(G)$是G的减函数。因而从效率的角度来看，圈地运动必然要求实行公共牧场的私有化，并按照羊头数的多少来划分土地专属权的大小。历史的发展恰恰与上述的理论分析是一致的，由于羊头数少的牧民占绝大多数，因而就出现了被马克思称之为“羊吃人”的现象。然而，如果没有圈地运动这场“农业革命”，随后的英国工业革命是不可能发生的。

“公共地悲剧”反映出了极其深刻的哲理——大家都拥有（公地）的自由，却导致了大家都受损的结局。这也正是哲学家怀海特把其称为“公共地悲剧”的原因。在怀海特的著作中这样写道：“悲剧的实质不在

于不幸，而在于(这种不幸是由大家共有)这一神圣性(所造成的)。"①要避免公共地悲剧的发生，如果排除私有化的可能性，靠技术是不可能解决的，它实际上取决于人们崇高的思想境界。

四、争议仲裁

在现实的经济活动中，经常会发生各种争议，例如劳资争议、合同争议、消费争议等。一旦争议发生，就需要社会的仲裁机构(通常为民事法庭)来加以仲裁。仲裁的主要形式有两种：协议仲裁和最后仲裁。根据民法原则，争议双方应首先自愿协商解决，只有争议双方无法达成协议时，仲裁机构才进行最后仲裁。这里我们引入一个工资最后仲裁模型，实际上这个模型可以推广到各式各样的争议仲裁博弈中。

假定参与争议的双方一为企业，一为工会，争议因工资而起。首先，企业和工会同时开出自己希望的工资水平，分别用 w_f 和 w_u 表示。如果 $w_f > w_u$，显然不存在争议。如果 $w_u > w_f$，双方就存在工资争议，并报仲裁机构进行仲裁。最后，仲裁机构在两者之中选择其一作为结果。这个博弈实际上属于我们后面要讲到的动态博弈，但由于仲裁机构超脱于该博弈，在这个博弈中不存在收益，所以我们可以把该博弈看作是静态博弈来处理，因为企业和工会报价之后，结果实际上已经决定了(不存在走后门，拉关系这样的可能)。在博弈中，我们把类似于仲裁机构的这种参与者称为自然，它参与博弈但却没有自己的利益。假定仲裁机构本身对工资水平有自己认为合理的值，用 x 来表示这一理想值。在企业和工会报价 w_f 和 w_u 之后，仲裁机构只是简单选择距 x 最为接近的要价：如果 $x < (w_f + w_u)/2$，仲裁者将选择 w_f；如果 $x > (w_f + w_u)/2$ 则选择 w_u，如果 $x = (w_f + w_u)/2$ 选择哪一个无关紧要，因为企业和工会工资报价的期望值相等。仲裁机构的仲裁原则如图 2.11 所示。而且仲裁机构遵循的这个仲裁原则是公共信息。对于 x，只有仲裁机构知道，企业和工会不知道。但 x 服从一定的概率分布是公共信息，其积累

① Whitehead, A. N. 1948: "Science and the Modern World", New York: Mentor, p17.

分布函数为 $F(x)$，概率密度函数为 $f(x)$。根据我们对仲裁机构的仲裁原则，如果双方的要价分别为 w_f 和 w_u，那么双方推断 w_f 被选中的概率为 $\text{Prob}\{w_f\}$ 和 w_u 被选中的概率为 $\text{Prob}\{w_u\}$，用数学表示出来为：

$$\text{Prob}\{w_f\}=\text{Prob}\{x<\frac{w_f+w_u}{2}\}=F(\frac{w_f+w_u}{2})$$

和 $$\text{Prob}\{w_u\}=1-F(\frac{w_f+w_u}{2})$$

据此，期望的工资水平为：

$$w_f\text{Prob}\{w_f\}+w_u\text{Prob}\{w_u\}$$
$$=w_fF(\frac{w_f+w_u}{2})+w_u[1-F(\frac{w_f+w_u}{2})]$$

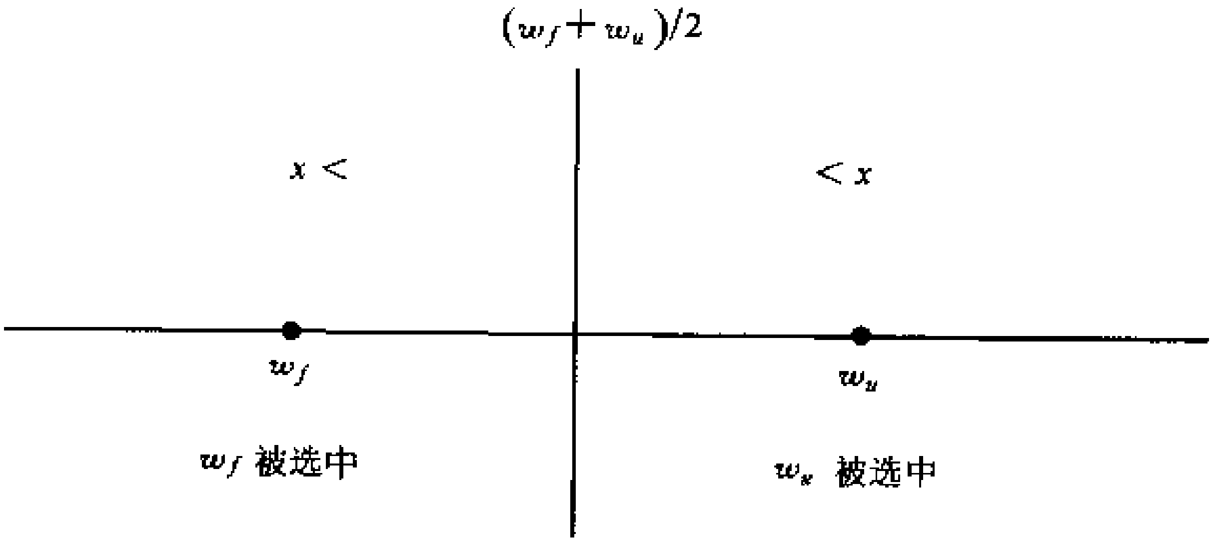


图 2.11　可能的仲裁结果

这里需要说明的是，$(w_f>)w_u$ # 相对于 $(w_f<)w_u$ 是工会的严格劣策略（你能回答为什么吗）。从企业的角度来说，w_f # $(>w_u)$ 相对于 $w_f(<w_u)$ 是企业的严格劣策略，因而从理论上而言不会出现 $w_f>w_u$ 的情况。

企业的目标是获得期望工资水平最小化的仲裁结果，工会则设法使其最大化。企业和工会之间形成了一种博弈关系，读者应当能够写出该博弈基本式的三个要素。

根据纳什均衡的定义，在给定工会策略 w_u 的条件下，企业追求支付最小化，即：

$$\min_{w_f} w_f F\left(\frac{w_f+w_u^*}{2}\right)+w_u^*\left[1-F\left(\frac{w_f+w_u^*}{2}\right)\right]$$

求一阶条件为：

$$\frac{1}{2}(w_u^*-w_f)f\left(\frac{w_f+w_u^*}{2}\right)=F\left(\frac{w_f+w_u^*}{2}\right) \tag{2.9}$$

在给定企业策略 w_f 的条件下，工会追求收益的最大化，即：

$$\max_{w_u} w_f^* F\left(\frac{w_f^*+w_u}{2}\right)+w_u\left[1-F\left(\frac{w_f^*+w_u}{2}\right)\right]$$

求一阶条件为：

$$\frac{1}{2}(w_u-w_f^*)f\left(\frac{w_f^*+w_u}{2}\right)=\left[1-F\left(\frac{w_f^*+w_u}{2}\right)\right] \tag{2.10}$$

如果纳什均衡为(w_f^*, w_u^*)，它一定同时满足(2.9)式和(2.10)式。这意味着(2.9)式和(2.10)式的左边项相等，因而有：

$$F\left(\frac{w_f^*+w_u^*}{2}\right)=\frac{1}{2} \tag{2.11}$$

这意味着企业和工会报价的平均值等于仲裁机构 x 分布的中值，这说明企业和工会都努力想与仲裁机构的偏好吻合。因为如果偏离 x 太远，那么对手的报价就越有可能被选中。将(2.11)带入(2.9)或(2.10)，得：

$$(w_u^*-w_f^*)=\frac{1}{f\left(\frac{w_f^*+w_u^*}{2}\right)} \tag{2.12}$$

工会和企业报价之差等于仲裁机构 x 分布的中值点的概率密度函数的倒数。为了更好地理解这一结果，我们假设 x 服从均值为 m，方差为 σ^2 的正态分布（也可以假设 x 服从均衡分布等），那么密度函数为：

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp[\frac{(x-m)^2}{2\sigma^2}] \tag{2.13}$$

根据(2.11)可得：

$$\frac{w_f{}^*+w_u{}^*}{2}=m \tag{2.14}$$

将(2.14)式带入(2.12)式，并根据(2.13)式得：

$$w_u{}^*-w_f{}^*=\frac{1}{f(m)}=\sqrt{2\pi}\sigma \tag{2.15}$$

根据(2.14)式和(2.15)式就能得到纳什均衡的要价为：

$$(w_u{}^*=m+\frac{\sqrt{2\pi}\sigma}{2},w_f{}^*=m-\frac{\sqrt{2\pi}\sigma}{2})$$

这表明双方的均衡要价以仲裁机构 x 的期望值 m 为中心对称，且要价之差随双方对仲裁机构 x 的不确定性(即 σ^2)的提高而增大。对这一均衡结果的直观解释非常简单，博弈的每一方都需进行权衡，一个更为不合理的要价(即工会更高的要价或企业更低的出价)一旦被仲裁机构选中就会给自己带来更高的收益，但其被选中的可能性却会相应降低。当企业和工会认为仲裁机构 x 的不确定程度增加(即 σ^2 变大)时，双方的要价就会更不合理。因为一个更不合理的要价与仲裁机构的 x 有较大差别的可能性变小了。相反，如果几乎不存在任何不确定性，双方都不敢开出一个离 x 期望值很远的要价，因为仲裁机构选择离 m 最近的报价的可能性非常大。

五、合作性谈判

我们考虑一个博弈模型：有一元钱，两兄弟分，两兄弟必须同时给出各自认为应得的份额。不妨假设兄弟 1 给出的份额为 x_1，而兄弟 2 给出的份额为 x_2。博弈的规则是，如果两兄弟给出的份额为 $x_1+x_2>1$，那么两兄弟只能得到 0；如果两兄弟给出的份额之和 $x_1+x_2\leqslant 1$，那么他们各自得到他们提出的份额。这实际上是一个简单的合作性博弈，合作

成功对双方都有利，而合作失败则双方都受损。根据前面对纳什均衡的定义，很容易知道这个合作性博弈存在无穷多个纳什均衡。凡是 $x_1, x_2 \geqslant 0$，并且 $x_1 + x_2 = 1$ 的所有 (x_1, x_2) 策略组合都是纳什均衡。因为如果给定兄弟 1 的策略 x_1，那么兄弟 2 的最优策略为 $1 - x_1$。因为如果 $x_2 > 1 - x_1$，那么双方都只能得到 0，显然这比得到 $1 - x_1$ 要差；如果 $x_2 < 1 - x_1$，显然也比得到 $1 - x_1$ 要差，因此只有 $x_2 = 1 - x_1$ 才是兄弟 2 的最优策略。对于兄弟 1 而言，给定兄弟 2 的策略 x_2，兄弟 1 的最优策略为 $x_1 = 1 - x_2$，分析同理，不再赘述。但我们在现实生活中，经常观察到的一个现象是，人们在分配利益时往往出现的均衡是平均分配，即兄弟 1 和兄弟 2 平均分配，各得 1/2 元钱。为什么会如此？这就引出了纳什最早开创的合作性谈判研究。

合作性谈判的模型是由纳什独立发展出来的，并由此开启了合作性博弈(Cooperative games)研究的大门。他把一个谈判问题极端地抽象化，除去所有的谈判细节和过程，而专注于讨论谈判的空间和谈判的结果之间的关系。具体讲，一个双人谈判问题，是两个人在所有谈判可能的结果所产生的效用空间里去协商，选出一个效用组合来。如果谈判不成，那么他们将各自得到一个谈判破裂的效用。纳什问的问题是，在不同的可能效用及协商破裂的可能下，怎样去选出一个公平的谈判结果。

定义 2.7　一个双人谈判问题(Two-person bargaining problem)，是一个 R^2 里的集合 X，以及 X 中的一个点 d，记作 (X, d)。其中 X 代表在所有谈判可能的结果下，谈判者的所有可能效用的集合。d 是谈判破裂点(Disagreement point)，代表当他们谈判破裂时，各自所得的效用。令 Σ 是所有谈判问题的集合。也就是说，$\Sigma = \{(X, d) \mid X \subset R^2, d \in X\}$。

我们假设 X 是凸集合(Convex set)和紧致集(Compact set)。再假设存在 $x \in X$，使得 $x_i > d_i, i = 1, 2$，即假设 X 中存在让两个人效用都大于谈判破裂时的效用，这样谈判才有意义。

定义 2.8　一个谈判问题的解，F 是一个从 Σ 映射到 R^2 的函数。也就是说，对于任意一个 $(X, d) \in \Sigma$，$F(X, d)$ 是一个 X 中的点。它代表的意思是，面对 (X, d) 这样一个谈判问题时，谈判者从最后的分配所得到的效用组合是 $F(X, d)$。

从定义 2.8 可以看出，纳什所找的不是一个单一的分配值，而是一个一般的、普遍的分配原则。在这个原则下，不论遇到什么分配问题，都可以用这个原则去分配。同时，纳什认为一个“好”的分配原则(即谈判的解)，应该满足下列四个公设：

A1. 效率性(Efficiency)

假如有一点 $x>F(X,d)$，则 $x\notin X$ 。这就是说谈判的解 F，必须是帕累托效率(Pareto efficiency)的分配。

A2. 尺度不变性(Invariance)

所谓不变性是指谈判的解与效用采用什么度量单位无关。也就是说，对于任何的 $u=(u_1,u_2)\in R^2$，令 $f_i(u_i)=a_iu_i+b_i$，且 $a_i>0$，其中 $i=1,2$。假若 u_i 代表第 i 个人的效用值，那么 $f_i(u)$等于是把第 i 个人的效用做一个线性变换。当我们把一个人的效用做线性变换时，他的效用其实并没有改变，只是用以衡量的单位变了。因此，他们的谈判解的惟一改变，也应该只是完全相同的单位改变而已。① 令 $f=(f_1,f_2)$，尺度不变性即是要求 $F(f(X),f(d))=f(F(X,d))$。这说明如果我们把原来问题的所有效用值都用 f 去变换，那么变换之后的问题的解，应该也是原先问题的解的相同变换。

A3. 无关选择的独立性(Independence of irrelevant alternative)

假设 $Y\subset X$，且 $F(X,d)\in Y$，则 $F(Y,d)=F(X,d)$，如图 2.12 所示。②这个公设的意思是说，假如在原先较大的谈判空间 X 下，$F(X,d)$ 选择了 Y 里的一点，那么表示 X/Y 个部分其实是无关紧要的。虽然 X 有大于 Y 的部分，但是 $F(X,d)$并不在 X/Y 里，而无关备选的独立性，即要求这一部分并不重要。因此，当它们被限制在较小的 Y 里而选择时，原先的解 $F(X,d)$这一点也是现在的解。

A4. 对称性(Symmetry)

如果原先的问题是完全对称的，那么两个人所得到的效用应该完全相等。也就是说，若 $(x_1,x_2)\in X\Rightarrow(x_2,x_1)\in X$，且 $d_1=d_2$，那么

---

① 微观经济学中偏好的是序数，而不是基数。

② 这个公设事实上并不是百分之百合理和毫无争议的。

$F_1(X,d)=F_2(X,d)$，如图 2.13 所示。

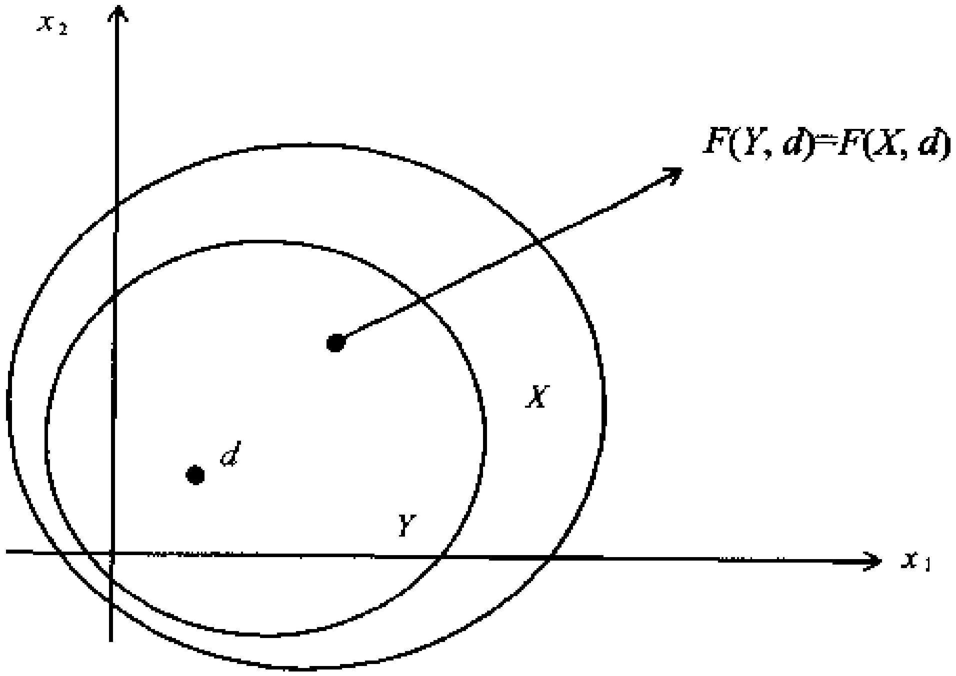


图 2.12　无关备选的独立性

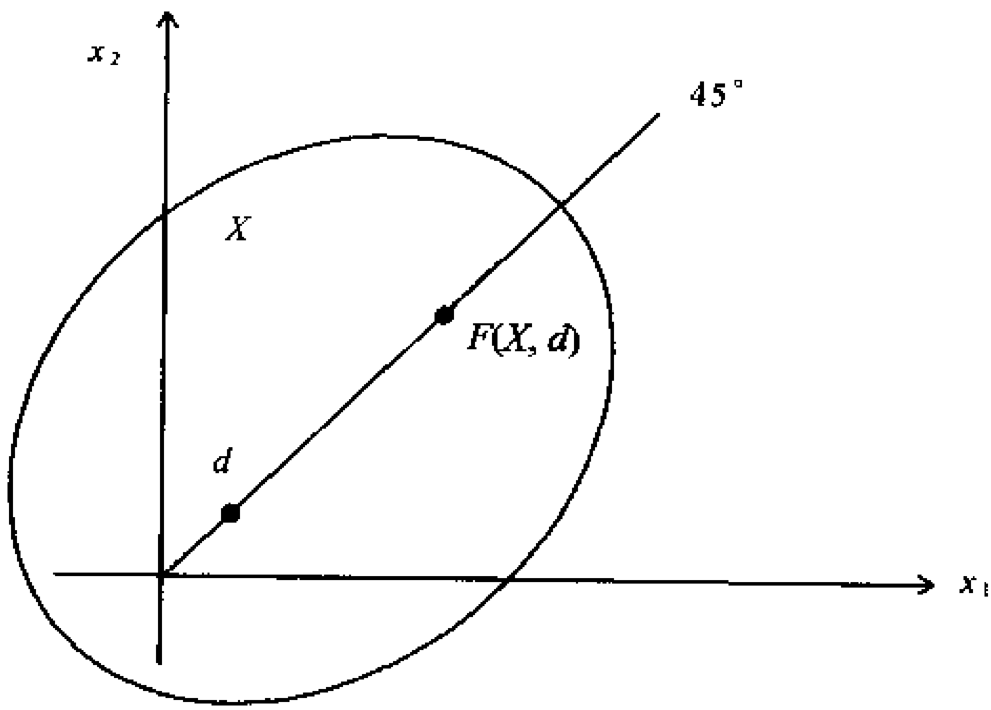


图 2.13　对称性

定理 2.1 (Nash)　上述双人谈判问题中只有一个解满足公设 **A1～A4**,那就是:

$$F(X,d)=\arg\max_{(x_1,x_2)\in X}(x_1-d_1)(x_2-d_2) \quad (a)$$

即 $F(X,d)$其实是以(d_1,d_2)为中心的双曲线与 X 的切点,如图 2.14 所示。

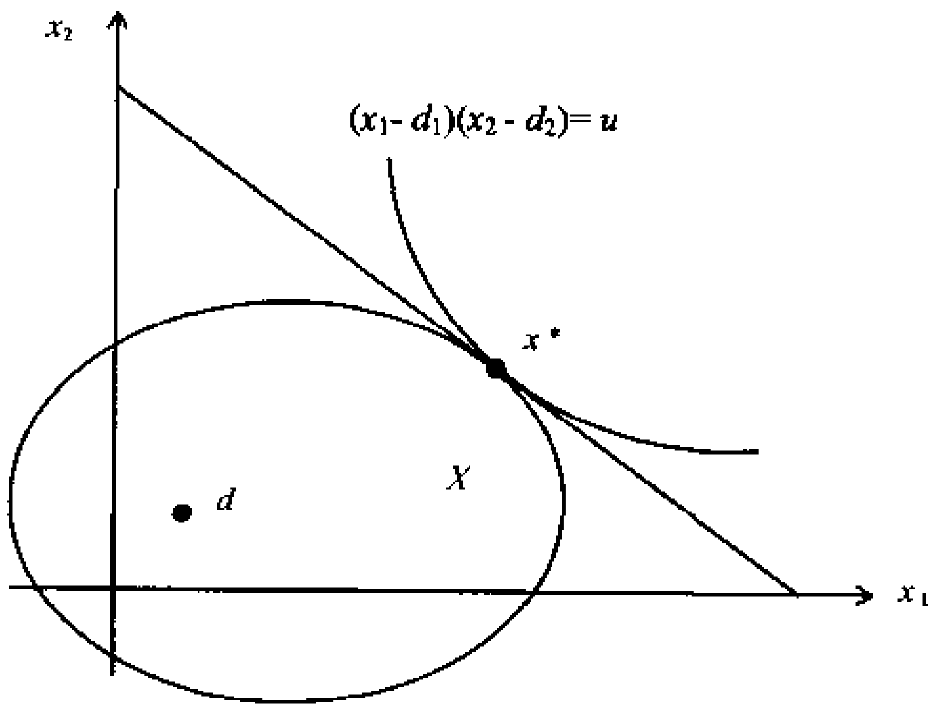


图 2.14　合作性谈判的解

证明:我们分两步来证明。

首先证明如果 $F(X,d)=\arg\max_{(x_1,x_2)\in X}(x_1-d_1)(x_2-d_2)$,则 F 一定满足公设 **A1～A4**。

(1)已知 $F(X,d)$是上述问题的最大化解。现在不妨假设存在一个 $y>F(X,d)$,并且 $y\in X$,那么可知 $F(X,d)$不是上述问题的最大化解,而这与已知条件 $F(X,d)$的最大化相矛盾。

(2)设 $f=ax+b,a>0$,则 $x^*\in F(f(X),f(d))=\arg\max[f_1(x_1)-f_1(d_1)][f_2(x_2)-f_2(d_2)]=a_1a_2F(X,d)$;同时 $x^*\in f(F(X,d))=\arg\max f[(x_1-d_1)(x_2-d_2)]=a\ F(X,d)+b$,显然对效用单位进行线性变换不会改变上述问题的分配方案。

(3)不妨设 $x^* \in F(X,d)$，已知 $Y \subset X$，并且 $x^* \in Y$，那么 x^* 一定是集合 Y 中的解，如果不是那就意味着存在一个 $y^* \in F(Y,d)$，有 $y^* > x^*$，而这意味着 x^* 不会是上述问题的最大化解($y^* \in X$)。显然与已知条件矛盾。

(4)显然 $(x_1-d)(x_2-d)=(x_2-d)(x_1-d)$。

较困难，是证明若有任何一个解 F 满足公设 **A1～A4**，那么 F 一定满足(a)式。令 $x^* = \arg \max\limits_{(x_1,x_2)\in X}(x_1-d_1)(x_2-d_2)$。换言之，$x^*$ 是 X 与(d_1,d_2)为中心的双曲线的切点。我们要证明如果任一个解 F 满足公设 **A1～A4**，那么 $F(X,d)=x^*$。

首先我们把问题标准化。也就是说，我们要找一个变换，把 x^* 的坐标转换成(1,1)，d 的坐标转换成(0,0)。令：

$$g_i(z)=\frac{z-d_i}{x_i^*-d_i}, i-1,2$$

则很容易得出 $g(x^*)=(1,1)$，且 $g(d)=(0,0)$。不但如此，原先和 X 相切于 x^* 点上的切线，在经过变换会成为一条和横轴成 45°，且和 X 相切于 $g(x^*)$点的线，如图 2.15 所示。

由于 X 是紧致集，因此 $g(X)$必然也是。同时，我们可以把 $g(x^*)$这一点上的切线延长，划出一对称于 45°线的等腰三角形$\triangle ABC$，使得 $g(X)\subset\triangle ABC$。我们先把 F 用在($\triangle ABC$,$g(d)$)这个谈判问题上。由于$\triangle ABC$ 以 45°线对称，且 $g_1(d)=g_2(d)=0$，因此由 **A3** 可知 $F(\triangle ABC,g(d))$一定在 45°线上。再由 **A1** 可知 $F(\triangle ABC,g(d))$是在这条线上满足帕累托效率的点。因此，$F(\triangle ABC,g(d))=(1,1)=g(x^*)$。由于 $g(X)\subset\triangle ABC$，而且 $F(\triangle ABC,g(d))=g(x^*)\in g(X)$，因此由 **A4** 可知 $F(g(X),g(d))=F(\triangle ABC,g(d))=(1,1)$。再令$f_i=g_i^{-1}$，也就是让 f_i 为 g_i 的反函数，则 $f_i(y)=(x_i^*-d_i)y+d_i$。很明显 f_i 也是一个 y 的线性转换。因此由 **A2** 可知 $F(X,d)=F(f(g(X)),f(g(d)))=f(F(g(X),g(d))=f(1,1)=x^*$。上式的第二个等式来自 **A2**。因此，我们证明了 $F(X,d)$一定等于 x^*。但 x^* 正是满足(a)式的解。

根据定理 2.1 可知，分钱博弈虽然有无穷多个纳什均衡，但只有惟

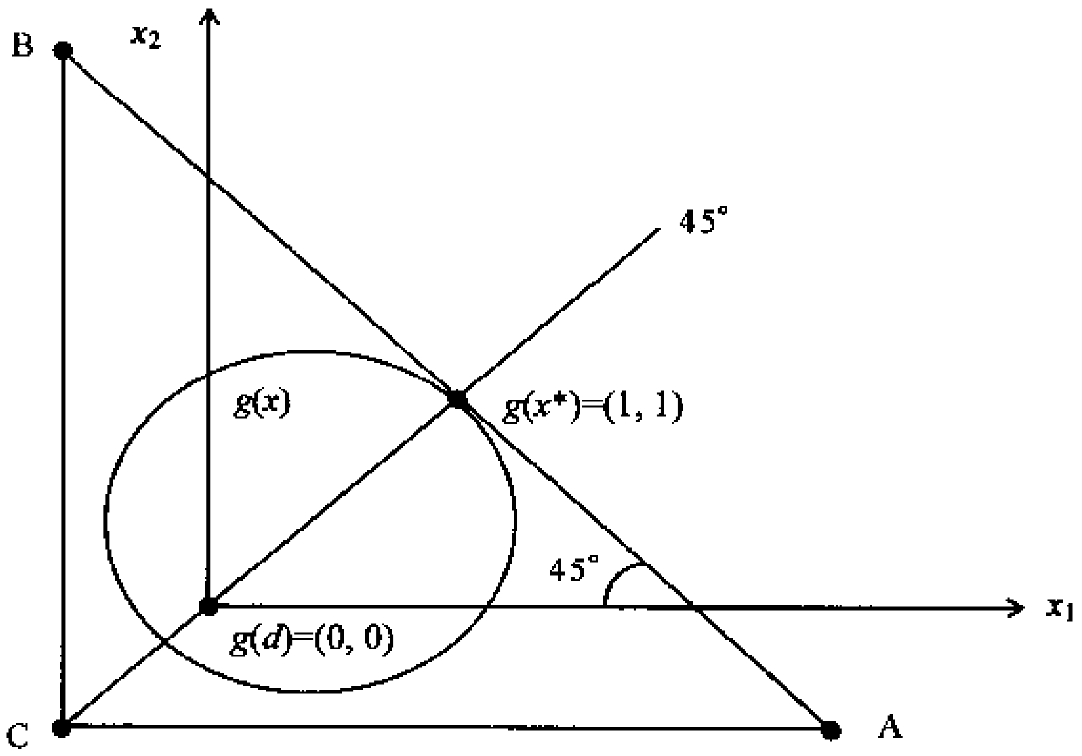


图 2.15 变换

——一个纳什均衡是公平的分配，即两兄弟各得 1/2 的钱。我们令 $d=(0,0)$，即如果两兄弟没有合作成功，那么他们得到 0 的收益，那么 $\{x_1=1/2,x_2=1/2\}$ 是 (a) 式惟一的最大化解。

定理 2.1 是一个非常优美、对称，且合乎情理与直觉的定理。它从逻辑的意义上证明了什么能称为公平（没有对错之分），揭示了为什么人们会认为平均分配就是公平。纳什对谈判问题的处理和证明方式，都深深地影响着后来合作博弈的发展。例如，他使用的公设法（Axiomatization），一直到今天，仍然是合作博弈研究者所使用的标准方法。

第四节 混合策略与纳什均衡的证明

由于纳什均衡是博弈论的基础，所以证明纳什均衡的存在性是必要的。但要证明纳什均衡的存在需要对策略进行一些扩展，即将纯策略

扩展为混合策略。前面讲到的策略实际上是纯策略，它是策略空间的基本元素。所谓基本就是我们不能对它再进行分解，而混合策略则至少能分解为两个（或以上）的纯策略。例如在石头、剪刀、布这个博弈中，孙悟空的一个混合策略为 1/3 的概率出石头、1/3 的概率出剪刀、1/3 的概率出布（目的是使猪八戒摸不着孙悟空的套路），它可以分解为三个纯策略　　石头、剪刀和布，显然石头（或剪刀或布）这个纯策略不可能再分解。按照前面给出的纳什均衡定义，许多博弈是不存在纳什均衡的，例如石头、剪刀、布这个博弈。但将纯策略扩展为混合策略，就能使纳什均衡这一概念适用于更为一般的情况。下面我们将首先给出混合策略的概念，然后证明纳什均衡的存在性。

一、混合策略

我们先来看石头、剪刀、布这个博弈，如图 2.16 所示。按照前面对纳什均衡的定义，显然该博弈不存在纳什均衡。因为每一个参与者都试图能先猜中对手可能采取的策略，从而选择相应的策略。例如，如果孙悟空猜测猪八戒可能采取石头的策略，那么孙悟空最优的策略是布；但是如果猪八戒猜测孙悟空可能采取布，猪八戒最优的策略是剪刀，而不再是石头。通俗地说，无论什么策略组合，孙悟空和猪八戒两者中总有一人希望改变策略，而不存在纳什均衡中所描述的参与者策略的一致性。

		猪八戒		
		石头	剪刀	布
	石头	0,0	+1,−1	−1,+1
孙悟空	剪刀	−1,+1	0,0	+1,−1
	布	+1,−1	−1,+1	0,0

图 2.16　石头、剪刀、布

类似于石头、剪刀、布这样的博弈实际上是非常普遍的，例如平常的扑克游戏、体育比赛、军事对抗等等。在这些博弈中，使诈是最有效的一种制胜手段，因为使诈能使对手猜不中自己要选择的策略，从而造成不确定性。参与者的最优行为是不确定的，面博弈的结果也必然包含这

种不确定性。军事中的“兵不厌诈”就形象地说明了这类博弈的要点。通过引入混合策略的概念，把纳什均衡的定义扩展到包含混合战略，就可以分析诸如石头、剪刀、布这类的博弈，例如扑克游戏、体育比赛、军事对抗等的均衡解。

我们用科学的方法表述，纯策略是最基本的策略，它不能够分解为更小的策略；而混合策略则是参与者策略空间 S_i 中全部纯策略的概率分布，它能够分解为更小的策略。也许把纯策略(Pure Strategy)译为基本策略，把混合策略(Mixed Strategy)译为复合策略更为合适些。对本章所分析的完全信息静态博弈来说，一个参与者的纯策略就是他可以选择的不同行动，例如在石头、剪刀、布博弈中，策略空间 S_i 内含有三个纯策略，分别为石头、剪刀和布，而参与者 i 的一个混合策略为概率分布 $(q,r,1-q-r)$，其中 q 为出石头的概率，r 为出剪刀的概率，$1-q-r\geqslant 0$ 为出布的概率，且 $0\leqslant q,r\leqslant 1$。由于纯策略可以看作是混合策略的一个特例，例如石头这个纯策略可以表述为混合策略 $(1,0,0)$，即以 1 的概率出石头，出剪刀和布的概率为 0，所以混合策略包含纯策略。

作为混合策略的第二个例子，请看图 2.17 所示博弈。在这个博弈中，有两位参与者，每个参与者有两个纯策略：硬币的正面和背面。如果两个参与者出硬币朝上的面一样，例如参与者 1 出硬币的正面，参与者 2 也出硬币的正面，那么参与者 1 赢得一个硬币；如果两个参与者出硬币朝上的面不一样，例如参与者 1 出硬币的正面，参与者 2 出硬币的反面，那么参与者 2 赢得一个硬币。这时参与者 1 的一个混合策略为概率分布 $(q,1-q)$，其中 q 表示出正面的概率，$1-q$ 表示出背面的概率，且 $0\leqslant q\leqslant 1$。在此博弈中，混合策略 $(1/2, 1/2)$ 表示参与者 1 出正面和背面的概率相同，而 $(1,0)$ 表示出正面的概率为 1，但绝不可能选择出背面。参与者 2 的情况和参与者 1 一样，不再赘述。参与者的纯策略只是混合策略的一种特例，混合策略 $(1,0)$ 表示的就是出正面这个纯策略。

		参与者2	
		正面	背面
参与者1	正面	+1,−1	−1,+1
	背面	−1,+1	+1,−1

图 2.17　猜币游戏

更为一般地，假设参与者 i 有 K 个纯策略 $S_i=\{s_{i1},\cdots,s_{iK}\}$，那么参与者的一个混合策略就是一个概率分布 $(p_{i1},\cdots,p_{iK})$，其中 $0\leqslant p_{ik}\leqslant 1$ 对所有 $k=1,\cdots,K$ 成立，且 $p_{i1}+\cdots+p_{iK}=1$。它表示参与者取策略 s_{i1} 的概率为 p_{i1}，取 s_{ik} 的概率为 p_{ik}，等等。以后我们用 p_i 来表示居于 S_i 的任意一个混合策略，用 P_i 来表示混合策略空间（所有混合策略的集合），用 s_i 来表示任意一个纯策略。

定义 2.9　对于博弈 $G=\{I,S,u\}$，假设 $S_i=\{s_{i1},\cdots,s_{iK}\}$，那么，参与者 i 的一个混合策略为映射 $p_i: S_i\rightarrow[0,1]$，即以 $p_i(s_{ik}\geqslant 0)$ 的概率取 S_i 中的纯策略 s_{ik}，其中 $\sum_{k=1}^{K}p_i(s_{ik})=1$。混合策略也可表示为 $p_i=(p_{i1},\cdots,p_{iK})$，其中对于所有的 $k=1,\cdots,K$，$0\leqslant p_{ik}\leqslant 1$，且 $p_{i1}+\cdots+p_{iK}=1$。

当参与者随机采取纯策略时，博弈结果的本身就是随机的（我们没有用不确定这个词），这时参与者的收益函数变为冯·诺伊曼—摩根斯坦恩（von Neumann-Morgenstern）收益函数，参与者 i 所得的收益是给定 n 个参与者混合策略组合 $p=(p_1,p_2,\cdots,p_n)$ 下的期望值。假设 $s=(s_1,\cdots,s_n)$ 为纯策略组合，其中 s_i 表示参与者 i 可用的任意一个纯策略，$S=S_1\times\cdots\times S_n$ 为策略空间的 n 重笛卡尔积。参与者 i 从混合策略组合 p 得到的 v—N—M（冯·诺伊曼—摩根斯坦恩）收益用数学可表示为：

$$v_i(p)=\sum_{s\in S}[p_1(s_1)p_2(s_2)\cdots p_n(s_n)]u_i(S) \qquad (*)$$

在前面分析严格劣策略的时候，很自然的一个结论是，如果一个策略是严格劣策略，那么参与者无论在什么情况下绝对不会选择它，但人们从不选择的策略是不是严格劣策略呢？在没有引人混合策略这个概

念前，这个逆命题是不成立的，而引入混合策略后这个逆命题成立。我们可通过图 2.18 来加以说明。

参与者 2

		左	右
参与者 1	上	3，—	0，—
	中	0，—	3，—
	下	1，—	1，—

图 2.18 虚构的博弈(1)

从图 2.18 中可以看出，在纯策略下参与者并没有严格劣策略，因为当参与者 2 的策略为右，参与者 1 选择下所得之收益要大于选上所得之收益(1>0)；而参与者 2 的策略为左时，参与者 1 选择下所得之收益要大于选中所得之收益(1>0)，因而不存在严格劣策略。但是，从这个博弈可以看出，下这个策略对于参与者 1 而言，是一个从来都不会选择的策略，因为当参与者 2 的策略为左时，参与者 1 最优的策略为上；当参与者 2 的策略为右时，参与者 1 最优的策略为中。从理论上而言，我们应当能够像剔除严格劣策略一样，剔除下这个策略。在引入混合策略之后，我们就会看到下是混合策略的严格劣策略。我们不妨假设，参与者 1 采取混合策略(p_{11}，p_{12}，0)，其中 p_{11}和 p_{12}大于 0，$p_{11}+p_{12}=1$。参与者 2 采取混合策略(q，$1-q$)，其中 $1\geqslant q\geqslant 0$。根据公式(*)参与者 1 的混合策略组合的期望值为：

$$E=(p_{11}\times q\times 3)+(p_{12}\times q\times 0)+[p_{11}\times(1-q)\times 0]+[p_{12}\times(1-q)\times 3]>1$$ ①

例如，我们取(p_{11}，p_{12}，0)=(1/2，1/2，0)，q 为[0，1]中任意值，参与者收益的期望值为 3/2>1，所以策略下相对混合策略(1/2，1/2，0)而言为严格劣策略，可以剔除。这里要记住只要某个纯策略严格劣于某个混合策略，它就是严格劣策略，可剔除。参与者无论在什么情况下从来都不会选择的策略一定是严格劣策略。这里要注意在纯策略下从来

① 通过一阶条件可知，期望值的最小值为 3/2。

都不选择的策略不一定为严格劣策略，例如图 2.19 所示。

		参与者 2 左	参与者 2 右
参与者 1	上	3,	0,—
	中	0,—	3,
	下	2,	2,—

图 2.19　虚构的博弈(2)

从图 2.19 可以看出，在纯策略的情况下，策略下是一个从不选择的策略，但它不是严格劣策略，即使引入混合策略，它仍不是严格劣策略，因为当参与者 2 的混合策略为 $(q, 1-q)$，$1/3<q<2/3$ 时，参与者 1 的混合策略 $(p_{11}, p_{12}, 0)$ 的期望值为 3/2，小于 2，所以下是参与者 1 的最优策略。之所以出现这种情况，原因在于它不满足“无论在什么情况下”这一条件。只有无论在什么情况下都不会被选择的策略才是严格劣策略。

二、纳什定理

我们分三个步骤来完成纳什定理的证明。第一步是将纳什均衡的概念扩展到混合策略，即混合策略纳什均衡；第二步求解石头、剪刀、布和猜币游戏这两个博弈的纳什均衡；第三步是一个“非数学化”的纳什均衡证明。

1. 混合策略纳什均衡

需要记住，纯策略只不过是混合策略的特例，也就是说，纯策略是混合策略的退化，纯策略集合是混合策略集合的真子集。

在纯策略下，纳什均衡的核心是每一个参与者所选择的策略均为对手策略的最优反应策略，并且所有参与者都如此。在混合策略下，同样应满足这一要求，纯策略要满足的条件必然是混合策略要满足的(因为纯策略是混合策略的真子集)。问题是混合策略纳什均衡是不是要比纯策略纳什均衡满足更多的条件。由于混合策略反映了参与者随机地选择策略，目的就是使对手摸不清自己的套路，因而博弈的结果也带有随机性，但是收益的期望值却是确定的(正是由于这一点，它才被看作

是完全信息，但又具有一点点非完全信息的因子)。

因而我们必须在两点上加以扩展：

(1)必须在混合策略的概念下来讨论最优反应策略，也就是任意一个人的最优反应策略是给定其他参与者混合策略组合下的混合策略。

(2)必须根据 v—N—M 收益函数来计算收益的期望值，因而所谓最优是期望值的最优。

如果博弈为 $G=\{I,S,u\}$，其中 $I=\{1,2,\cdots,n\}$，$S_i=\{s_{i1},\cdots,s_{iK}\}$，$u=(u_1,u_2,\cdots,u_n)$，$\forall\, i\in I$，那么参与者 i 的混合策略为 $p_i=(p_{i1},p_{i2},\cdots,p_{iK})$，期望收益函数为 $v_i(p_i,p_{-i})=E[u_i(s)]$。

定义 2.10 一个混合策略组合 $p^*=(p_1^*,p_2^*,\cdots,p_n^*)$是博弈 $G=\{I,S,u\}$的一个纳什均衡，如果对每一个参与者 $i=1,2,\cdots,n$，对于所有的 $p_i\in P_i$，不等式 $v_i(p_i^*,p_{-i})\geqslant v_i(p_i,p_{-i})$都成立。

显然，混合策略的纳什均衡与纯策略下的纳什均衡并没有本质的区别，关键就在于它是根据 v—N—M 收益函数得到参与者的期望值。

2. 两个例子

第一个例子为猜币游戏(见图 2.17)。这个博弈显然不存在纯策略下的纳什均衡，但却存在着混合策略纳什均衡。将该博弈转化为基本式：

(1)参与者：游戏者 1 定义为 1，游戏者 2 定义为 2，$I=\{1,2\}$。

(2)混合策略空间：参与者 i 的策略空间为 $S_i=\{$正面，背面$\}$，s_i 为纯策略，混合策略为 $p_i=(p_{i1},1-p_{i1})$，$1\geqslant p_{i1}\geqslant 0$ 为参与者 i 出正面的概率，$1-p_{i1}\geqslant 0$ 为参与者 i 出背面的概率。混合策略空间为 $P_i=\{p_i: 1\geqslant p_{i1}\geqslant 0; 1-p_{i1}\geqslant 0\}$。

(3)收益函数：参与者 i 的收益函数为 $u_i(s_1,s_2)$，v—N—M 收益函数为 $v_i(p_1,p_2)=\sum\limits_{s\in S}[p_1(s_1)p_2(s_2)]u_i(s_1,s_2)$。

给定游戏者 2 的混合策略为$(p_{21},1-p_{21})$，那么游戏者 1 出正面的期望值为：

$$p_{21}\times 1+(1-p_{21})\times -1 \tag{2.16}$$

游戏者 1 出背面的期望值为：

$$p_{21}\times -1+(1-p_{21})\times 1 \tag{2.17}$$

如果游戏者 1 的混合策略为$(p_{11},1-p_{11})$，那么它的期望收益为：

$$v_1(p_1,p_2)=p_{11}\times(2.16)+(1-p_{11})\times(2.17)$$

整理后得：

$$\max\ v_1(p_1,p_2)=(2p_{11}-1)(2p_{21}-1)$$

当 $p_{21}>1/2$ 时，$\frac{dv_1}{dp_{11}}>0$，这说明增加 p_{11}能够增加收益。它的意思非常简单，如果游戏者 2 出正面的概率大于 1/2(出正面的可能性大于背面的可能性)，游戏者 1 最优的策略反应是增加出正面的概率，当$p_{11}=1$时，期望收益最大。如果游戏者 2 出正面的概率小于 1/2(出背面的可能性较大)，游戏者 1 最优的策略反应是增加出背面的概率，当$p_{11}=0$时，期望收益最大。显然，像我们在前面讲到的参与者的策略是相互依存的，游戏者 1 的 p_{11}是 p_{21}的函数，①就像古诺模型中的策略函数。

$$p_{11}(p_{21})=\begin{cases}0 & p_{21}<1/2\\ [0,1] & p_{21}=1/2\\ 1 & p_{21}>1/2\end{cases} \tag{2.18}$$

(2.18)式我们称为策略对应，因为当 $p_{21}=1/2$ 时，p_{11}为集合[0,1]。如图 2.20 所示。

给定游戏者 1 的混合策略为$(p_{11},1-p_{11})$，那么游戏者 2 出正面的期望值为：

$$p_{11}\times -1+(1-p_{11})\times 1 \tag{2.19}$$

游戏者 2 出背面的期望值为：

$$p_{11}\times 1+(1-p_{11})\times -1 \tag{2.20}$$

如果游戏者 2 的混合策略为$(p_{21},1-p_{21})$，那么它的期望收益为：

① 确切地说，是对应。函数是点到点的映射，而对应是点到集合的映射。

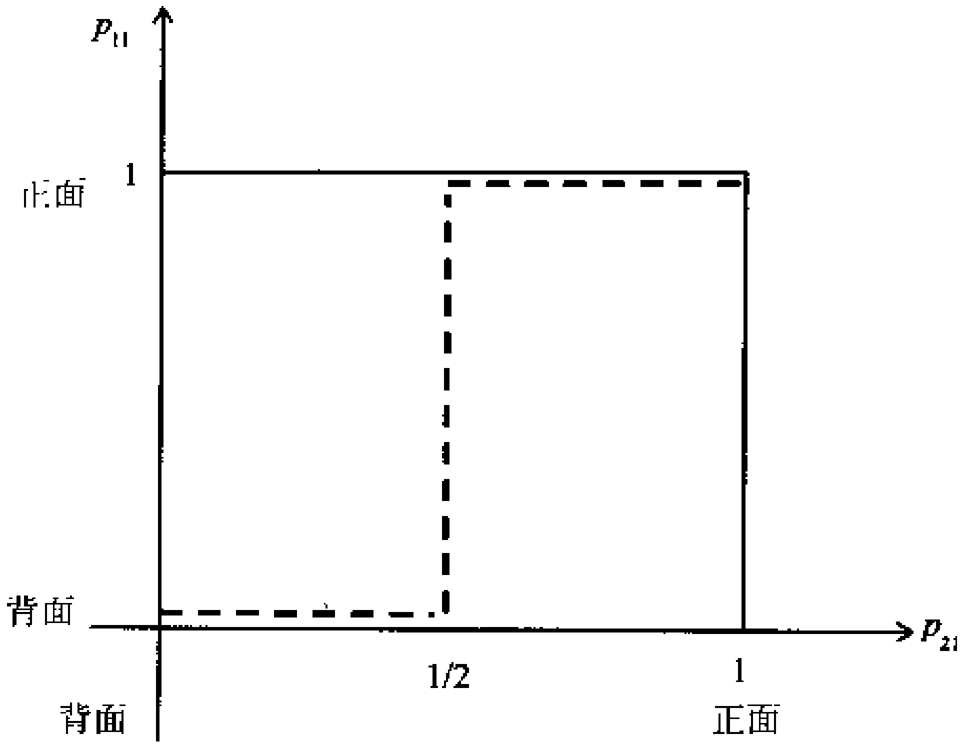


图 2.20　游戏者 1 的策略对应

$$v_2(p_1,p_2)=p_{21}\times(2.19)+(1-p_{21})\times(2.20)$$

整理后得：

$$\max v_2(p_1,p_2)=(1-2p_{21})(2p_{11}-1)$$

当 $p_{11}<1/2$ 时，$\frac{\mathrm{d}v_2}{\mathrm{d}p_{21}}>0$，这说明增加 p_{21}能够增加收益。它的意思非常简单，如果游戏者 1 出正面的概率小于 1/2，游戏者 2 最优的策略反应是增加出正面的概率，当 $p_{21}=1$ 时，期望收益最大。如果游戏者 1 出正面的概率大于 1/2，游戏者 2 最优的策略反应是增加出背面的概率，当 $p_{21}=0$ 时，期望收益最大。游戏者 2 的 p_{21}是 p_{11}的对应。

$$p_{21}(p_{11})=\begin{cases}1 & p_{11}<1/2\\ [0,1] & p_{11}=1/2\\ 0 & p_{11}>1/2\end{cases} \tag{2.21}$$

(2.21)式如图 2.21 所示。

如果混合策略组合(p_1^*,p_2^*)为纳什均衡，那么必须同时满足(2.18)

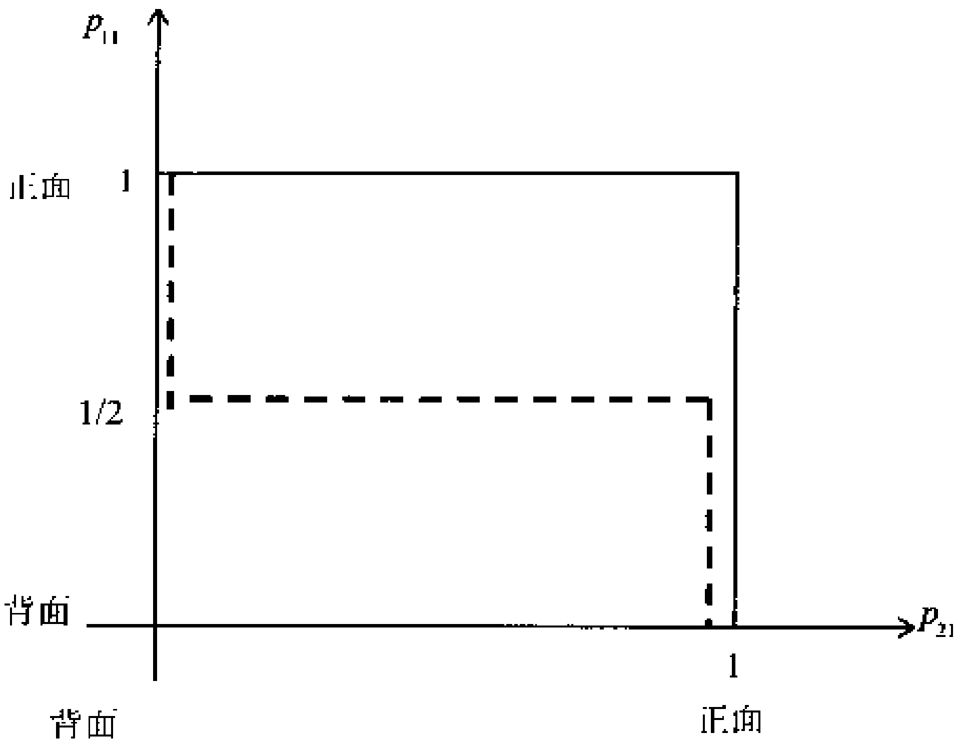


图 2.21　游戏者 2 的策略对应

和(2.21)两式。纳什均衡$(p_1^*, p_2^*)=((1/2,1/2),(1/2,1/2))$,即游戏者 1 出正面和背面的概率各为 1/2,游戏者 2 出正面和背面的概率各为 1/2。如在图 2.22 中,只有当 $p_{11}=p_{21}=1/2$ 时,游戏者的策略对应曲线才有相交点,这个交点就是纳什均衡。

第二个例子是石头、剪刀、布(见图 2.16)。它比猜币博弈复杂得多。首先将这个博弈转化为基本式:

(1)参与者:孙悟空定义为 1,猪八戒定义为 2,$I=\{1,2\}$。

(2)混合策略空间:参与者 i 的策略空间为 $S_i=\{$石头,剪刀,布$\}$,s_i 为纯策略。混合策略为 $p_i=(p_{i1}, p_{i2}, 1-p_{i1}-p_{i2})$,$p_{i1}$为参与者 i 出石头的概率,p_{i2}为参与者 i 出剪刀的概率,$1-p_{i1}-p_{i2}$为出布的概率,并且满足非负与总和为 1 的条件。混合策略空间为 $P_i=\{p_i:\geqslant p_{i1}, p_{i2}\geqslant 0; 1-p_{i1}-p_{i2}\geqslant 0\}, i=1,2$。

(3)收益函数:博弈参与者 i 的收益函数为 $u_i(s_1, s_2), i=1,2$。参与者 1 的 v—N—M 收益函数为 $v_1(p_1, p_2)=\sum_{j=1}^{3} p_{1j}\sum_{k=1}^{3} p_{2k}u_1(s_{1j}, s_{2k})$,参

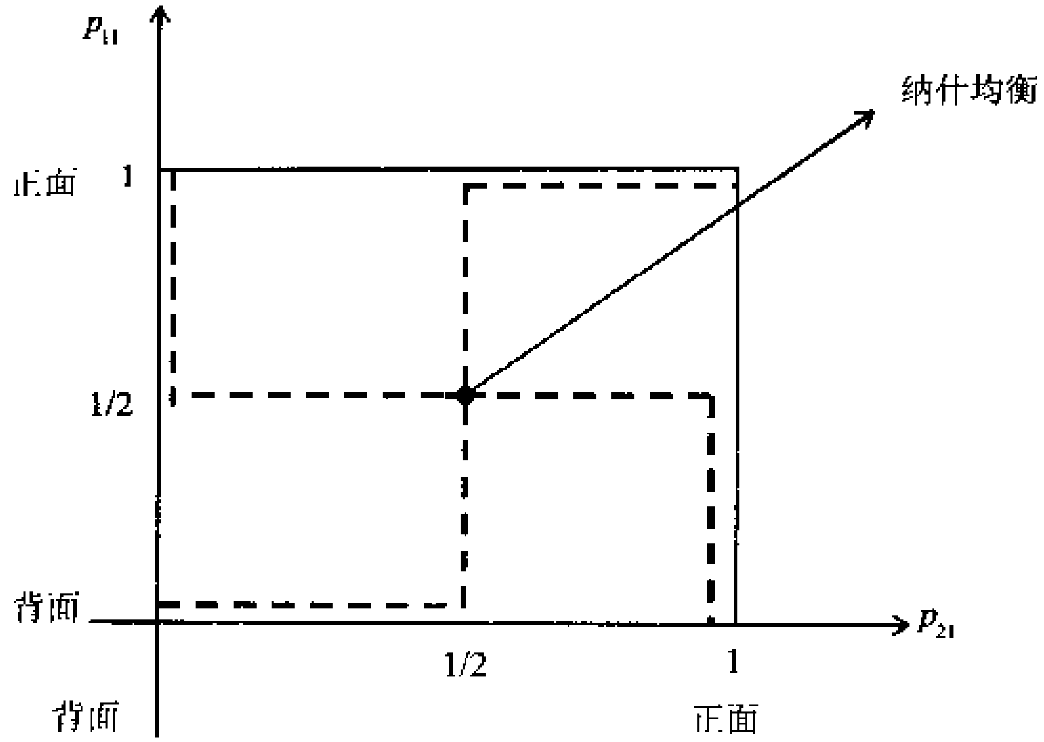


图 2.22 纳什均衡点

与者 2 的 v—N—M 收益函数为 $v_2(p_1,p_2)=\sum_{k=1}^{3}p_{1k}\sum_{j=1}^{3}p_{2j}u_1(s_{1j},s_{2k})$。

给定猪八戒的混合策略为 $(p_{21},p_{22},1-p_{21}-p_{22})$，孙悟空出石头的期望值为：

$$p_{21}\times 0+p_{22}\times 1+(1-p_{21}-p_{22})\times -1 \tag{2.22}$$

出剪刀的期望值为：

$$p_{21}\times -1+p_{22}\times 0+(1-p_{21}-p_{22})\times 1 \tag{2.23}$$

出布的期望值为：

$$p_{21}\times 1+p_{22}\times -1+(1-p_{21}-p_{22})\times 0 \tag{2.24}$$

如果孙悟空的混合策略为 $(p_{11},p_{12},1-p_{11}-p_{12})$，那么他的期望收益为：

$$v_1(p_1,p_2)=p_{11}\times(2.22)+p_{12}\times(2.23)+(1-p_{11}-p_{12})\times(2.24)$$

整理之后得：

$$\max v_1(p_1,p_2)=p_{11}(3p_{22}-1)-p_{12}(3p_{21}-1)+p_{21}-p_{22} \tag{\#1}$$

对上式求一阶偏导数，得：

$$\frac{\mathrm{d}v_1}{\mathrm{d}p_{11}}=3p_{22}-1=0 \qquad \frac{\mathrm{d}v_1}{\mathrm{d}p_{12}}=1-3p_{21}=0$$

当 $p_{22}>1/3$ 时，$\frac{\mathrm{d}v_1}{\mathrm{d}p_{11}}>0$，这说明 p_{11} 增大能够增加收益；当 $p_{21}<1/3$时，$\frac{\mathrm{d}v_1}{\mathrm{d}p_{12}}>0$，这说明 p_{12}增大能够增加收益。通俗地说，如果猪八戒出剪刀的概率大于 1/3，那么孙悟空增加出石头的概率(p_{11})能够增加收益；如果猪八戒出石头的概率小于 1/3，那么孙悟空多出剪刀能够增加收益。这说明如果猪八戒的混合策略不为(1/3，1/3，1/3)，那么孙悟空就会相应地增加或是减少出石头或剪刀或布的概率。

现在给定孙悟空的混合策略为(p_{11}，p_{12}，$1-p_{11}-p_{12}$)，那么猪八戒出石头的期望值为：

$$p_{11}\times 0+p_{12}\times 1+(1-p_{11}-p_{12})\times -1 \tag{2.25}$$

出剪刀的期望值为：

$$p_{11}\times -1+p_{12}\times 0+(1-p_{11}-p_{12})\times 1 \tag{2.26}$$

出布的期望值为：

$$p_{11}\times 1+p_{12}\times -1+(1-p_{11}-p_{12})\times 0 \tag{2.27}$$

如果猪八戒的混合策略为(p_{21}，p_{22}，$1-p_{21}-p_{22}$)，那么它的期望收益为：

$$v_2(p_1,p_2)=p_{21}\times(2.25)+p_{22}\times(2.26)+(1-p_{21}-p_{22})\times(2.27)$$

整理之后得：

$$\max\ v_2(p_1,p_2)=p_{21}(3p_{12}-1)-p_{22}(3p_{11}-1)+p_{11}-p_{12} \tag{\#2}$$

对上式求一阶导数，得：

$$\frac{dv_2}{dp_{21}}=3p_{12}-1=0 \qquad \frac{dv_2}{dp_{22}}=1-3p_{11}=0$$

当 $p_{12}>1/3$ 时，$\frac{dv_2}{dp_{21}}>0$，这说明 p_{21} 增加能够增加收益；当 $p_{11}<1/3$时，$\frac{dv_2}{dp_{22}}>0$，这说明 p_{22}增加能够增加收益。通俗地说，如果孙悟空出剪刀的概率(p_{12})大于 1/3，那么猪八戒增加出石头的概率(p_{21})能够增加收益；如果孙悟空出石头的概率小于 1/3，那么猪八戒多出剪刀能够增加收益。这说明如果孙悟空的混合策略不为(1/3,1/3,1/3)，那么猪八戒就会相应地增加或是减少出石头或剪刀或布的概率。

综上所述，如果猪八戒的混合策略为(1/3,1/3,1/3)，那么孙悟空最优的混合策略为(1/3,1/3,1/3)；反之，如果孙悟空的混合策略为(1/3,1/3,1/3)，那么猪八戒最优的策略为(1/3,1/3,1/3)，所以混合策略组合((1/3,1/3,1/3),(1/3,1/3,1/3))为纳什均衡。

3. 纳什定理

前面我们分析了各种各样的完全信息静态博弈，并且都能求出这些博弈的纳什均衡解，是不是所有的博弈都存在纳什均衡呢？或者说在什么情况下我们能够肯定地知道存在纳什均衡？显然这个问题至关重要。对这个问题的回答最早(1950 年)是由纳什给出，被称为纳什定理。

定理 2.2 纳什定理　如果完全信息静态博弈 $G=\{I,S,u\}$是有限的，即参与者是有限的，S 包含的纯策略也是有限的，那么一定存在纳什均衡。

我们分析的博弈都满足这两个条件，参与者有限——通常为两个人；S 也是有限的——要么存在若干个纯策略，例如石头、剪刀、布，要么虽然纯策略有无穷多，但却是有界闭集，例如古诺模型中的 $S_i=[0,a]$，所以前面的博弈都存在纳什均衡。

对纳什定理的证明我们分几个步骤来完成：第一步，介绍不动点定理；第二步，说明纳什均衡就是不动点。第三步，完成证明。

第一步，不动点定理。

我们先来看图 2.23。

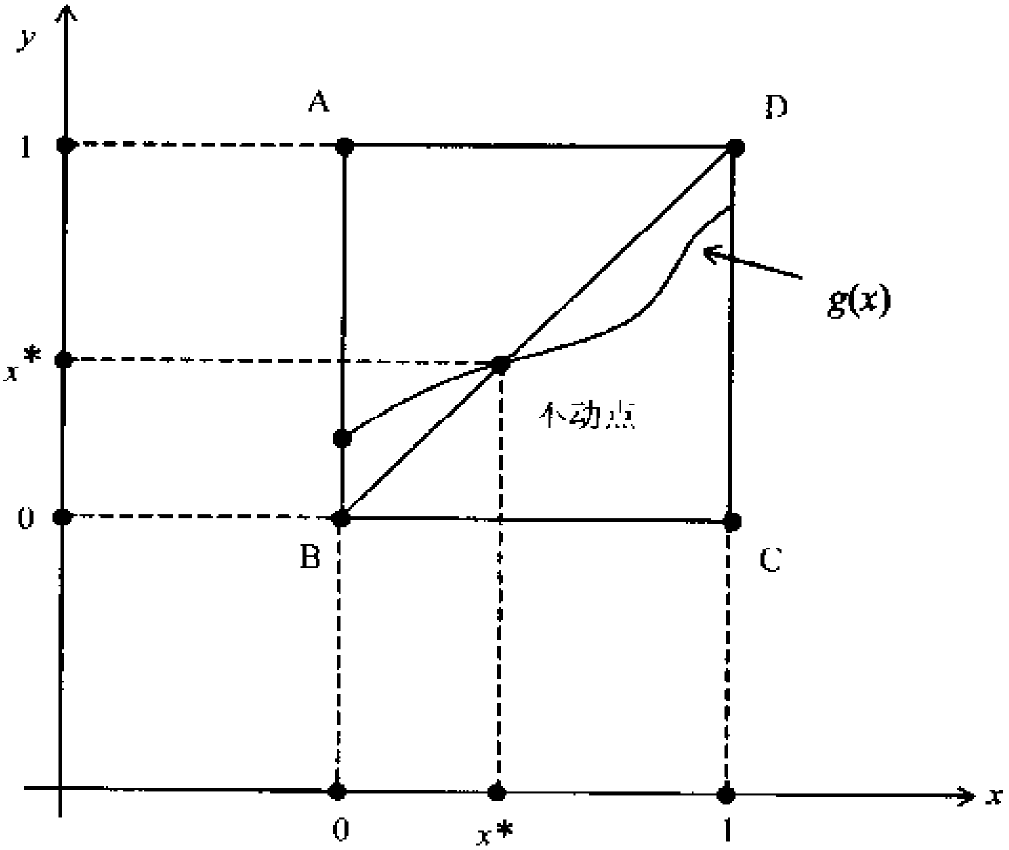


图 2.23　从区间[0,1]到区间[0,1]的连续函数具有不动点

图 2.23 中的方框为正方形。BD 为 45 度线，它实际上等价于函数 $y=f(x)=x$。BD 线上的每一点都满足 $f(x)=x$，我们把满足 $f(x)=x$ 的 x 称为不动点。BD 线所代表的函数的自变量 x 取值为[0,1]，称为定义域，而因变量 y 的取值为[0,1]，称为值域。定义域和值域都为[0,1]的函数 $f(x)$，又可写作 $f:[0,1]\rightarrow[0,1]$，读作从[0,1]到[0,1]的映射。现在我们做一个小小的游戏。你拿一支笔，从 AB 线出发任意画一条曲线，记住你可以任意画，但必须连续画不能断裂，直到 CD 线结束。这时你会发现无论你如何画，这条曲线一定会和 BD 线相交，这个交点就是不动点。因为你任意画的曲线等价于一个函数 $g(x)$，没有断裂意味着它是连续的，与 BD 线相交意味着 $g(x^*)=x^*$，所以交点为不动点。

所有任意画的曲线都有如下的要点：

(1)定义域[0,1]是非空的、有界的、闭的、凸集合，简称非空凸紧集。有界闭集即为紧集。

(2)函数 $f(x)$ 是连续的，并且是自身对自身的映射。

上面我们考虑的只是一元函数，对于 n 元函数是不是如果满足上述要点就存在不动点呢？数学家布劳尔(Brower)证明了这一点。

定理 2.3 布劳尔(Brower)不动点定理　假设 $A\subset R^n$ 是一个非空凸紧集合，并且 $f: A\rightarrow A$ 是一个连续函数，那么 $f(\cdot)$ 必有一个不动点，即存在一个 $x\in A$ 使得 $x=f(x)$。

这里 R 为实数，R^n 为实数的 n 重笛卡尔积，x 为 n 维向量。在前面的例子中 $n=1$，$A=[0,1]$。在实际运用中，我们常常遇到的不是函数而是对应的情况，例如猜币游戏这个博弈存在的是策略对应(2.18)和(2.21)，而不是策略函数，这就有了对布劳尔不动点定理的推广。

定理 2.4 角谷(Kakutani)不动点定理　假设 $A\subset R^N$ 是非空凸紧集，并且 $f: A\rightarrow A$ 是上半连续对应，并且对于每一个 $x\in A$，集合 $f(x)\subset A$ 是非空凸紧集，那么 $f(\cdot)$ 必有一个不动点，即存在一个 $x\in A$ 使得 $x\in f(x)$。

角谷不动点定理与布劳尔不动点定理的区别在于：

(1)映射 f 为对应而不是函数，所以 $f(x)$ 为一个集合(一个点也能组成集合，所以函数是对应的特例)。

(2)映射 f 是上半连续的。一个对应只有既是上半连续又是下半连续才能称为连续，因而上半连续比连续的要求要弱。

第二步，纳什均衡是不动点。

根据纳什均衡的定义，在给定其他参与者策略的情况下，参与者 i 选择一个策略以使其收益最大化，并且对每一个参与者都如此，即纳什均衡是一个使每一个参与者收益最大的策略组合。其他参与者的策略不同，参与者 i 相应采取的策略也不同，这在博弈论中称为策略依存，在数学上称为策略函数或策略对应。如果把参与者 i 的策略函数(对应)定义为 $R_i(p_1,\cdots,p_{i-1},p_{i+1},\cdots,p_n)=p_i$，其中 $i=1,\cdots,n$，那么 n 个参与者的策略函数(对应)就构成了一个方程组，我们定义为 $R(p_1,\cdots,p_n)$，即：

$$R(p_1,\cdots,p_n)=\begin{cases}R_1(p_2,\cdots,p_n)=p_1\\ \vdots\\ R_i(p_1,\cdots,p_{i-1},p_{i+1},\cdots,p_n)=p_i\\ \vdots\\ R_n(p_1,\cdots,p_{n-1})=p_n\end{cases}$$

如果$(p_1^*,\cdots,p_n^*)$是纳什均衡，它一定满足$R(p_1^*,\cdots,p_n^*)=(p_1^*,\cdots,p_n^*)$，即纳什均衡是一个不动点。例如，前面分析的古诺模型中的策略函数(2.1)和(2.2)共同构成了一个方程组$R(q_1,q_2)$，这个方程组的解就是不动点。

第三步，完成证明。

现在我们要完成的证明是：(1)策略函数R_i的定义域为非空凸紧集；(2)R_i为连续函数或是上半连续的对应。由于R是R_i的方程组，R_i满足这两个条件，其中$i=1,\cdots,n$，总的策略方程组R必然满足这两个条件，而根据不动点定理我们知道存在着不动点$(p_1^*,\cdots,p_n^*)$，使得$R(p_1^*,\cdots,p_n^*)=(p_1^*,\cdots,p_n^*)$。

(1)定义域为非空凸紧集。如果博弈有n个参与者，参与者i有k个纯策略，那么策略方程组R的定义域为$[0,1]^{k\times n}$，它表示$[0,1]$的$k\times n$重的笛卡尔积。例如函数$f(x,y)$，$x,y\in A$，那么函数的定义域就为$A^{1\times 1}$，即x，y都从集合A中取值。由于混合策略的取值为$[0,1]$，而我们知道$[0,1]$为非空凸紧集，所以$[0,1]^{k\times n}$也是非空凸紧集。在古诺模型中，它的定义域虽然与我们上面讲的有所不同，即纯策略有无穷多个，但策略q(产量)的取值为$[0,a]$，它同样是非空凸紧集。这表明证明纳什均衡真正需要的并不是“有限的”要求，而是策略空间必须为紧致集合。如果一个参与者有两个纯策略x和y，那么所有混合策略集合(策略空间)就是满足$p_x+p_y=1$的直线，其中p_x和p_y都大于等于零。如果有三个纯策略x、y和z，那么所有混合策略集合(策略空间)就是一个满足$p_x+p_y+p_z=1$的一个三角平面，其中p_x、p_y和p_z都大于等于零。如图2.24所示。

所谓有界集合就是我们任意画一个圆(圆的半径可以任意大，但不

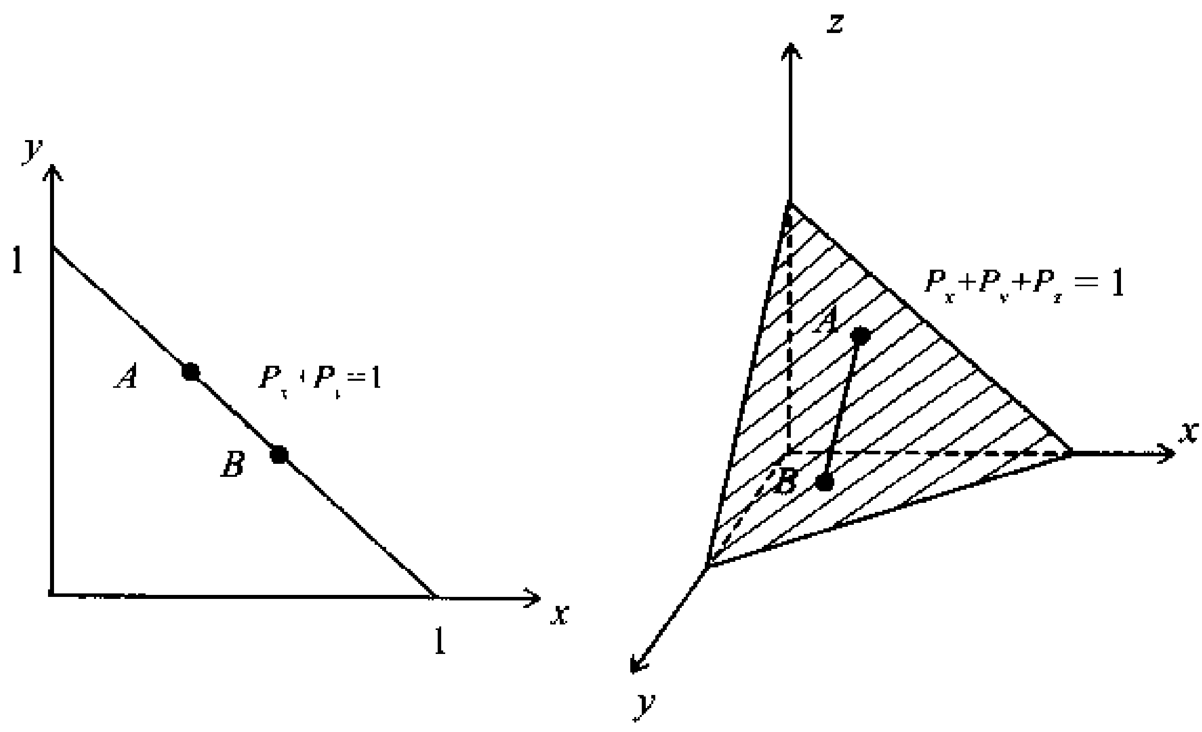


图 2.24 非空凸紧集

能无穷大)可以把它包下。所谓闭集就是包含边界的集合。所谓凸集就是在集合内任意找两点,连接两点的直线一定位于集合中。图 2.24 中的直线和三角平面显然是有界的,同时也是闭的,即它们是紧致集合,简称紧集;同时它们也是凸集。

(2)策略函数是连续的或策略对应是上半连续的。要证明策略函数是连续的或策略对应是上半连续的,需要很好的数学功底,特别是拓扑学的知识,这里不做证明。一般而言,从混合策略来理解策略,策略函数就是连续的,至少是上半连续的。

总而言之,策略方程组 $R(p_1,\cdots,p_n)$满足不动点定理的两个条件,因而存在着不动点,而不动点就是纳什均衡,这就完成了纳什定理的证明。需要指出的是,纳什定理只是表明纳什均衡存在的充分条件,而不是必要条件,因而不满足以上这两个条件的博弈也有可能存在纳什均衡。纳什定理只表明存在纳什均衡,而不排除存在着多个纳什均衡的可能,而多重均衡会使得我们无法有效地对博弈的结果作出预测。

复习思考题

1. 什么是纳什均衡？为什么说纳什均衡构成了博弈论的基础？纳什均衡的意义是什么？

2. 请写出争议仲裁的基本式的三个要素，并根据课文中的已知条件分析三党政治的纳什均衡。

3. 见下图烛光晚餐。请你规范地求出该博弈的纳什均衡。

		情侣 2 利顺德	情侣 2 喜来登
情侣 1	利顺德	+1,+1	−1,−1
情侣 1	喜来登	−1,−1	+1,+1

烛光晚餐

经济学系列教材

第三章

非完全信息静态博弈

在这一章中，我们主要讨论非完全信息静态博弈。非完全信息是指博弈中至少有一位参与者的收益函数不是公共信息，或者说参与者的收益是不确定的。在博弈论中，非完全信息博弈又称为贝叶斯博弈(Bayesian Game)。贝叶斯博弈都是动态博弈，因为通过豪尔绍尼转换，可以把非完全信息静态博弈转换为动态博弈，它是一种特殊的动态博弈。实际上所有的静态博弈都可以转化为动态博弈，至于是用动态的观点还是用静态的观点，主要取决于分析是否方便。用基本式来分析静态博弈要方便一些，不过这是相对的而不是绝对的，特别是包含复杂的非完全信息博弈更是如此，即使该博弈是静态的。

我们首先会从一个例子开始，说明非完全信息博弈的性质，从而引出豪尔绍尼转换

(Harsanyi Transformation)的概念；其次，给出静态贝叶斯博弈的基本式和贝叶斯纳什均衡的概念。静态贝叶斯博弈主要用于拍卖分析和拍卖机制设计，因而拍卖博弈和显示偏好原理就成为本章最后的内容。

第一节 豪尔绍尼转换

我们首先来看一个例子，这个例子我们称为囚徒困境 2。

例 3.1 囚徒困境 2。

我们假设囚徒之间存在江湖道义，囚徒有讲道义和不讲道义之分。两个囚徒被警察抓获，囚徒 1 是一个讲道义的囚徒，这是公共信息；但囚徒 2 是否讲道义，囚徒 1 却不清楚，他有可能讲道义，也有可能不讲道义，囚徒 2 知道自己是否讲道义，即囚徒 2 是否讲道义是一个私人信息。这个非完全信息(不对称信息)博弈可用图 3.1 表示。

		囚徒 2(讲道义) S	囚徒 2(讲道义) C		囚徒 2(不讲道义) S	囚徒 2(不讲道义) C
囚徒 1	S	−1, −1	−9, −5	S	−1, −1	−9, 0
	C	−5, −9	−6, −6	C	−5, −9	−6, −6

图 3.1 囚徒困境 2

S 表示沉默，C 表示招供。如果囚徒是讲江湖道义的囚徒，他招供而另一个囚徒没有招供，那么他会受到良心的谴责，给他带来 −5 的效用；如果囚徒是不讲江湖道义的囚徒，即使他出卖朋友他也不会受到什么良心的谴责，带给他的效用为 0；如果两个囚徒都招供，那么谁也不欠谁。其他的内容与原来的囚徒困境一样。

在囚徒困境 2 这个博弈中，相当于存在着两个博弈。如果囚徒 1 相

信因徒 2 是一个讲道义的人，他们的纳什均衡为(S,S)和(C,C)。[①] 如果因徒 1 相信因徒 2 是一个不讲道义的人，那么纳什均衡为(C,C)。到底哪一个策略组合是整个博弈的纳什均衡显然无法回答。其原因首先在于因徒 1 对因徒 2 不同的信念(Belief)导致了即使是相同的策略，也会出现不同的收益函数和不同的收益，即在相同的策略组合下，收益具有不确定性。例如，(S,C)这个策略组合，对因徒 2 而言就有不同的收益，在因徒 2 讲道义时，收益为－5，在因徒 2 不讲道义时，收益为 0。这一点显然是所有非完全信息博弈都具有的共同特性，这就是为什么从这一点来定义非完全信息的原因。其次，信念不同出现的均衡解也会不同。例如，在因徒 1 相信因徒 2 的情况下，均衡解为(S,S)，而在因徒 1 不相信因徒 2 的情况下，均衡解为(C,C)。但什么情况下相信，什么情况下不相信，并不存在标准。

由于参与者的收益函数具有不确定性，因而不可能通过求解最大化的方式找到最优策略，换句话说，就是什么策略都可能成为最优策略，任何结果都有可能是博弈的均衡解。这实际上意味着我们对贝叶斯博弈根本没办法分析。如果博弈不是像因徒困境 2 那么简单，而是稍微复杂一些，情况就更是如此。

因徒困境 2 还可以用扩展式表示出来。

图 3.2 所表示的因徒困境 2 与图 3.1 所表示的因徒困境 2 完全等价。为了简单一些，我们省掉了收益。图 3.2 实际上清楚地表明了因徒困境 2 所面临的困难，该博弈有两个开始点，在因徒 1 行动的时候，因徒 1 分不清他到底位于哪一个节点，是 x_3，x_4，x_5，还是 x_6。我们把参与者分不清的节点集合称为信息集(Information Set)，如果博弈的扩展式中的每一个信息集都只包含一个节点，那么我们就称这个博弈是完美信息的(Perfect Information)。在图 3.2 中，博弈的扩展式有三个信息集，它们分别是 $h_1=\{x_1\}$，$h_2=\{x_2\}$ 和 $h_3=\{x_3,x_4,x_5,x_6\}$。显然完美信息一定是完全信息，但反之不成立，这说明完美信息要比完全信息严格。由于该博弈有两个开始点，可以理解为两个不同的博弈，但关键是，

① 均衡(S,S)>均衡(C,C)，所以(S,S)是一个较为合理的纳什均衡。

这两个博弈被一条虚线连接起来，因而它又是一个博弈。它既是两个博弈又是一个博弈，从逻辑上来说这是矛盾的，因而我们不可能直接分析它。两个开始点还意味着从博弈的结束点往回推，我们不知道会回到哪一个开始点，这实际上意味着如果参与者要达到某个结束点，那么不知道采取什么样的策略才能达到这个结束点，也就是说没有人能保证一定走到这个结束点。因而均衡会不会出现只有上帝才知道了。当然最优策略也就无从谈起。

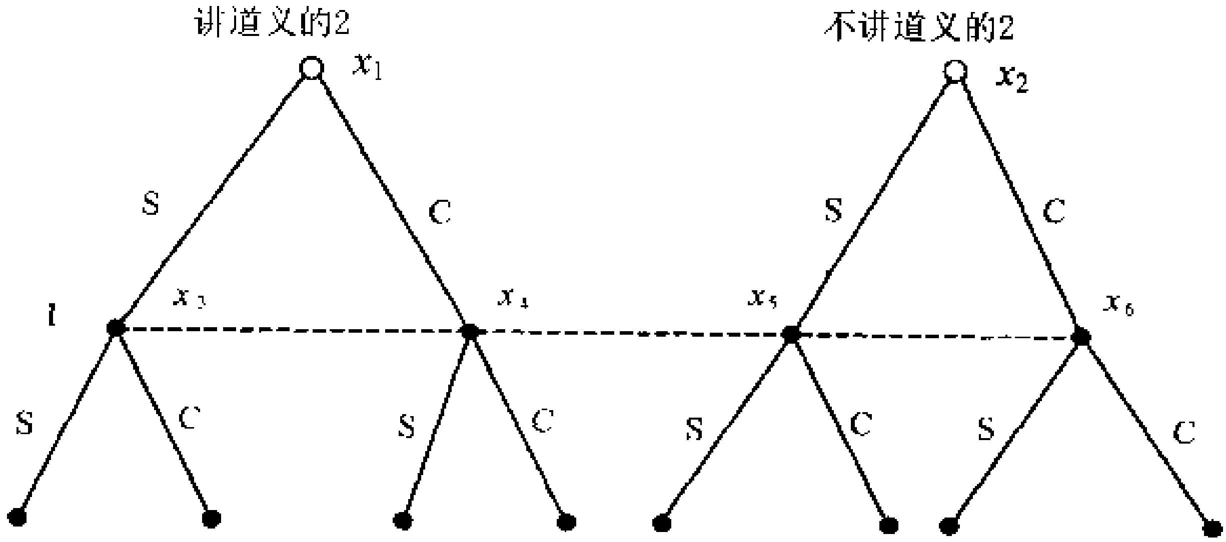


图 3.2　囚徒困境 2

面对这种两个开始点的困难，一个好的想法就是引入概率分布，即囚徒 2 为讲道义的人的概率为 θ，不讲道义的概率为 $1-\theta$，这样一处理，就意味着囚徒 1 可以计算出如果他采取某种策略，那么他可能实现的期望值。通过比较不同的期望值，就能得出作为一个“理性人”囚徒 1 应选择的策略，而囚徒 2 知道囚徒 1 会按照期望值的大小来决定相应的策略，因而囚徒 2 相应也知道他应采取什么样的策略。比如，尽管囚徒 2 是一个讲道义的人，但在知道囚徒 1 认为他是个不讲道义的人的可能性较大（θ 非常小）的情况下，囚徒 1 招供的期望值一定会大于沉默的期望值，因而囚徒 1 的最优策略是招供，在这种情况下，囚徒 2 最优的策略只能是招供，尽管囚徒 2 是一个讲道义的人。

把囚徒 2 是什么样的人看作一种概率分布，显然是一个非常妙的想法，妙就妙在它把一个既是两个博弈又是一个博弈这样的矛盾化解

了(变为一个博弈),并能够利用 v—N—M 期望效用原则来分析参与者的收益(期望值),从而使最优策略乃至纳什均衡都有了意义。最早把这一想法变成规范性的学术语言是由经济学家豪尔绍尼在 1967 年完成的,[①]也正因为这个贡献,豪尔绍尼与纳什和塞尔腾(Selten)分享了 1994 年的诺贝尔经济学奖。

我们把参与者的性质定义为类型(Type),用 t 来表示,T 为类型的集合,称为类型空间。以囚徒困境 2 为例,t_{11}表示囚徒 1 是讲道义的人,t_{21} 表示囚徒 2 是讲道义的人,t_{22}表示囚徒 2 是不讲道义的人,$T_1=\{t_{11}\}$,$T_2=\{t_{21},t_{22}\}$。

定义 3.1 豪尔绍尼转换——自然(Nature)决定参与者类型的概率分布,并且这个概率分布是公共信息。

豪尔绍尼转换是一种思路。我们也可以这样来理解豪尔绍尼转换:囚徒 1 通过抛硬币来决定囚徒 2 的类型,如果硬币正面朝上,那么囚徒 1 就相信囚徒 2 是讲道义的人;如果硬币背面朝下,那么囚徒 1 就相信囚徒 2 是不讲道义的人,因而囚徒 2 的类型(t_{21}=讲道义,t_{22}=不讲道义)的概率分布就为(1/2,1/2)。如果囚徒 1 认为囚徒 2 是讲道义的人的概率为 1/10000,那么囚徒 1 可以同时抛 14 个硬币,如果 14 个硬币都正面朝上($1/2^{14}\approx1/10000$),那么囚徒 1 就相信囚徒 2 讲道义。

有了豪尔绍尼转换,图 3.2 中的扩展式可转变为图 3.3 中的扩展式。

从图 3.3 可以看出,自然首先行动,决定囚徒 2 类型的概率分布,囚徒 2 知道他的类型,但是囚徒 1 不知道,所以囚徒 1 的信息集 $h_1=\{x_1,x_2\}$,即囚徒 1 不能区分他到底位于节点 x_1 还是 x_2。囚徒 2 的信息集为 $h_2=\{x_3,x_4\}$和 $h_3=\{x_5,x_6\}$,表明这是一个静态博弈。

图 3.2 中的扩展式还可以转变为图 3.4 中的扩展式。

图 3.3 和图 3.4 清楚地说明了豪尔绍尼转换的意义,它克服了非完全信息博弈既是两个(多个)博弈又是一个博弈这样的矛盾,并且把非完全信息博弈转换成一个非完美信息博弈来处理。

---

① Harsanyi, J. 1967: "Games with Incomplete Information Played by Bayesian Players", *Management Science* 14, pp. 159~182, 320~334, 486~502.

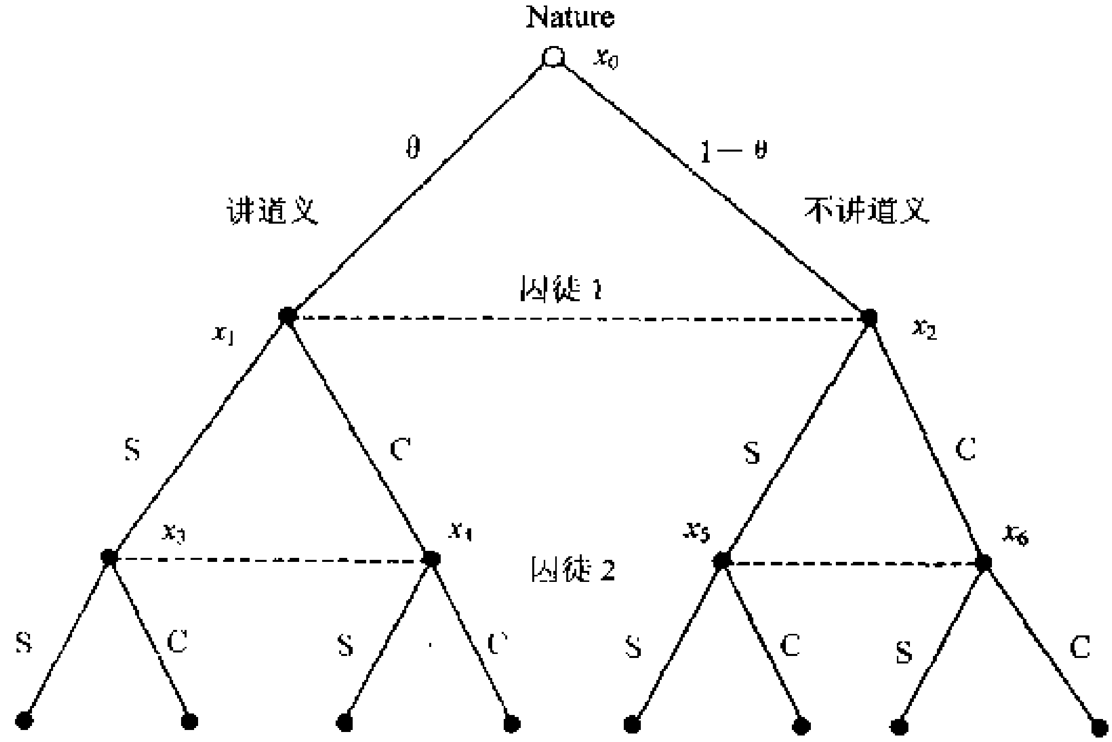


图 3.3 囚徒困境的豪尔绍尼转换

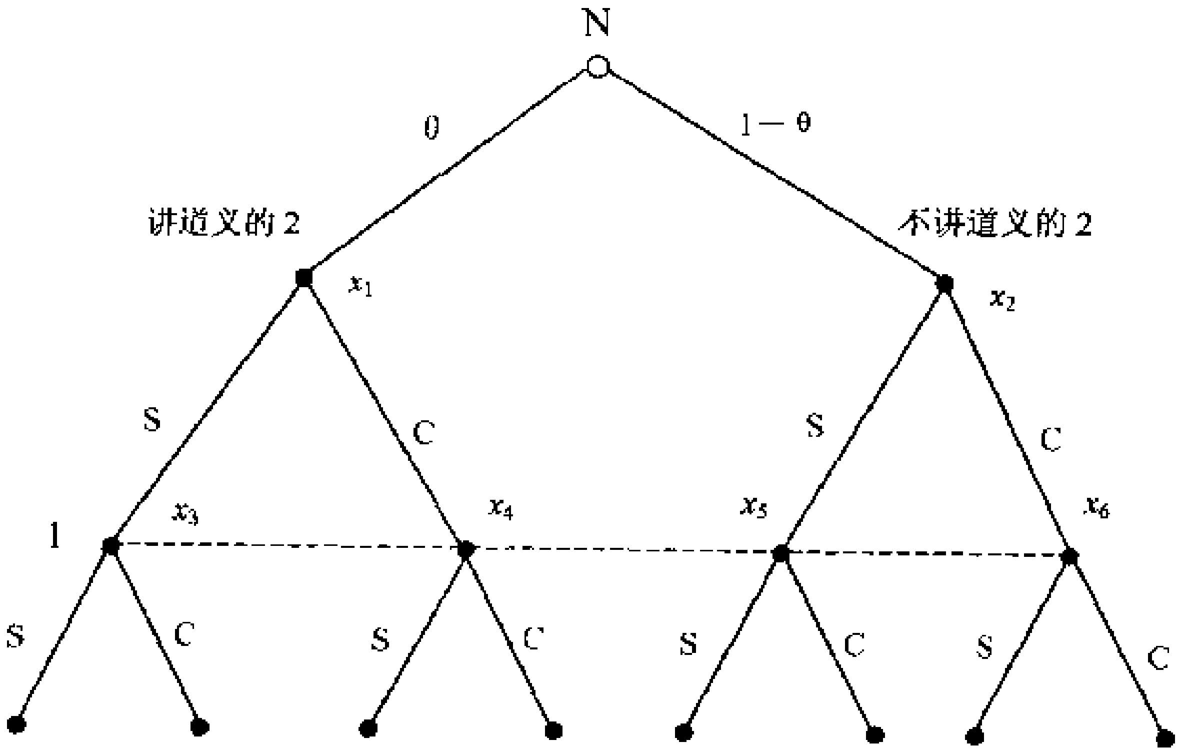


图 3.4 囚徒困境的豪尔绍尼转换

第二节　贝叶斯纳什均衡

在完全信息静态博弈中，参与者的策略等于行动；而在贝叶斯静态博弈中，参与者的策略是类型的函数，类型不同采取的行动也就不同。这说明贝叶斯静态博弈的基本式与完全信息静态博弈的基本式有所不同，需要做一些处理。在这一节中，首先给出贝叶斯静态博弈的基本表达式，在此基础上给出贝叶斯纳什均衡。

一、贝叶斯静态博弈的基本式

第二章中定义的 n 人完全信息静态博弈的基本式为 $G=\{I,S,u\}$，其中 $S=\{S_1,\cdots,S_n\}$，S_i 为参与者 i 的策略空间；$u=(u_1,\cdots,u_n)$，$u_i(s_1,\cdots,s_n)$ 为所有参与者分别选择策略 $(s_1,\cdots,s_n)$ 时参与者 i 的收益。由于完全信息静态博弈中的策略实际上等于一个简单的行动，因而基本式又可写作 $G=\{I,A,u\}$，其中 $A=\{A_1,\cdots,A_n\}$，A_i 为参与者 i 的行动空间；$u=(u_1,\cdots,u_n)$，$u_i(a_1,\cdots,a_n)$ 为所有参与者分别选择行动 $(a_1,\cdots,a_n)$ 时参与者 i 的收益。在贝叶斯静态博弈中，博弈的基本式显然不能简单地表示为这种形式。首先，策略实际上是依赖于参与者的类型，策略的实质是参与者根据自身的类型来选择相应的行动，策略不再是简单的行动。其次，收益函数也成为类型的函数，类型不同参与者的收益函数自然不同，如果收益函数与参与者的类型无关，就意味着贝叶斯博弈中的收益函数退化为完全信息静态博弈中的收益函数，完全信息博弈又可看作参与者的类型始终如一，即不存在信息不完全的情况。显然博弈不再是贝叶斯博弈而是完全信息博弈。若能指出完全信息博弈是贝叶斯博弈的特例这一点，就能理解上述内容。

要建立非完全信息静态博弈的基本式表述，首先要表示出非完全信息的关键因素，即每一参与者知道他自己的收益函数，但不能确知其他参与者的收益函数。令参与者 i 可能的收益函数表示为 $u_i(a_1(t_1),$

$\cdots, a_i, \cdots, a_n(t_n); t_i)$，其中 t_i 称为参与者 i 的类型，它属于一个可能的类型空间 T_i，每一类型 t_i 都对应着参与者 i 不同的收益函数的可能情况。如果按上述方法在参与者的收益函数和类型之间建立一种对应关系，那么参与者知道自己的收益函数就等于说参与者知道自己的类型，参与者 i 不能确定其他参与者的收益函数就等于说参与者 i 不知道其他参与者的类型。我们令 $t=(t_1,\cdots,t_n)$，$T=(T_1,\cdots,T_n)$，$t_{-i}=(t_1,\cdots,t_{i-1},t_{i+1},\cdots,t_n)$，用 T_{-i} 表示 t_{-i} 所有可能的集合。用概率 $p(t_{-i}|t_i)$ 来表示参与者 i 在知道自己的类型的前提下，对其他参与者类型的联合分布（即 t_{-i}）概率的信念。根据豪尔绍尼转换，联合分布 $p(t)$ 和 $p(t_i)$ 是公共信息，因而参与者 i 会根据贝叶斯法则来计算概率 $p_i(t_{-i}|t_i)$。在统计学中，经常会假设参与者之间的类型相互独立，在这种情况下，$p(t)=p(t_1)\times\cdots\times p(t_n)$，$p(t_{-i}|t_i)=p(t_1)\times\cdots\times p(t_{i-1})\times p(t_{i+1})\times\cdots\times p(t_n)$。为了使我们的说明更形象一些，引入一个改造后的古诺模型。

贝叶斯古诺模型：所有的内容和第二章提到的内容一样，不一样的地方在于：(1)参与者为三个垄断企业，定义为 1、2、3；(2)每一个企业的成本有可能为高成本 c_H 和低成本 c_L，第一个企业为低成本的概率为 θ_1，高成本的概率为 $1-\theta_1$；第二个企业、第三个企业依此类推，即 θ_2 和 $(1-\theta_2)$、θ_3 和 $(1-\theta_3)$。为了简单起见，假设企业的类型彼此独立。如果每个企业知道自己的类型，但不能确定其他企业的类型，那么 $T_1=T_2=T_3=\{c_L,c_H\}$，$t_{i1}=c_L$，$t_{i2}=c_H$。所有的类型组合 t 有八种情况，即 $\{(c_L,c_L,c_L),(c_L,c_L,c_H),\cdots,(c_H,c_H,c_H)\}$。联合概率分布 $p(t_{11},t_{21},t_{31})=p(c_L,c_L,c_L)=\theta_1\times\theta_2\times\theta_3$，$p(t_{21},t_{31}|t_{11})=\theta_2\times\theta_3$，$p(t_{11})=\theta_1$。产量 q_i 就是参与者 i 的行动，行动空间 $A_i=[0,a]$，参与者 1 的收益函数 $u_1(q_1(t_1),q_2(t_2),q_3(t_3))$，其中 $q_i(t_i)$ 就是策略，收益函数 u_i 是一个从 $A\times T$ 到 R（实数）的映射。读者可以自己写出其他的概率分布和收益函数，只要依此类推就行。

定义 3.2 一个贝叶斯静态博弈的基本表达式可记作 $G_B=\{I,A,p,T,u\}$。其中 I 是所有参与者的集合；$A=(A_1,\cdots,A_n)$，A_i 是参与者 i 的行动空间；$T=(T_1,\cdots,T_n)$，T_i 为参与者 i 的类型空间；p 是定义在

T 上的概率联合分布函数，它代表各参与者对其他参与者可能类型的先验信念；$u=(u_1,\cdots,u_n)$，u_i 代表参与者 i 的收益，它是一个从 $A\times T$ 到 R 的映射。也就是说，参与者的收益是他们所用的行动的函数，也是参与者类型的函数。

对于参与者的信念 p，有几点需要重点说明：

第一，概率 $p(t)$ 和 $p(t_i)$ 由自然决定，并成为公共信息。但它又可以被理解成是参与者的一种先验概率。这种先验概率不是凭空产生，而是来自于(理性的)参与者对自然的所有的了解后最科学地得出的。它既是客观的，又是主观的。比如，中央电视台每晚的天气预报，它是利用最科学的手段(即充分理性)根据所有的能够获得的(自然)信息，对明天下雨的概率作出的主观猜测，由于它是现有技术条件下所能得到的天气情况最好的预测，所以它又是客观的，并被所有的电视观众所知(公共信息)。

第二，贝叶斯静态博弈 G_B 实际上是一个基本式博弈 G，只不过每一个不同的类型 $t\in T$ 对应着一个不同的基本式博弈。这实际上意味着我们求贝叶斯博弈只不过是在计算一串基本式博弈的期望值，豪尔绍尼转换的意义就在于此。

第三，概率 $p(t_{-i}|t_i)$ 表示参与者 i 在知道自己的类型的前提下(其实是否知道自己的类型并不重要)，对其他参与者类型的联合分布(即 t_{-i})概率的信念。它实际上是参与者 i 根据贝叶斯法则(Bayes Rule)计算得到的后验概率。对于其他参与者而言，他们知道参与者 i 对他们各自类型所持有的推断，反之也一样，即所有 $p(t_{-i}|t_i)$ 是一个共同信息，之所以如此，在于 $p(t)$ 和 $p(t_i)$ 是共同信息。

如果 t 是连续的，那么：

$$p(t_{-i}^{*}|t_i)=\frac{p(t_{-i}^{*},t_i)}{\int_{t_{-i}\in T_{-i}}p(t_{-i},t_i)\mathrm{d}t_{-i}}=\frac{p(t_{-i}^{*},t_i)}{p(t_i)}$$

如果 t 是离散的，那么：

$$p(t_{-i}^* \mid t_i) = \frac{p(t_{-i}^*, t_i)}{\sum_{t_{-i} \in T_{-i}} p(t_{-i}, t_i)} = \frac{p(t_{-i}^*, t_i)}{p(t_i)}$$

由于策略是一个从类型到行动的映射，通过行动和类型可以构建出所有的策略，所以参与者的策略空间没有在贝叶斯静态博弈的基本表达式 G_B 中出现。

二、贝叶斯纳什均衡

纳什均衡是一个策略组合，要定义贝叶斯纳什均衡，首先必须定义贝叶斯静态博弈中参与者的策略。

定义 3.3　在贝叶斯静态博弈 $G_B=\{I,A,T,p,u\}$ 中，参与者 i 的一个纯策略 s_i 是一个由 T_i 映射到 A_i 的函数。对于 T_i 中的每一个 t_i，$s_i(t_i)$ 表明自然赋予参与者 i 的类型为 t_i 时，i 将从行动空间 A_i 中选择一个行动。

给出了贝叶斯静态博弈关于策略的定义之后，我们就可以定义贝叶斯纳什均衡。

定义 3.4　在贝叶斯静态博弈 $G_B=\{I,A,T,p,u\}$ 中，策略组合 $s^*=(s_1^*,\cdots,s_n^*)$ 是一个纯策略贝叶斯纳什均衡，如果对每一个参与者 i 及对 i 的类型空间 T_i 中的每一个类型 t_i，$s_i^*(t_i)$ 满足：

$$\max_{a_i \in A_i} \sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) u_i(s_1^*(t_1),\cdots,s_i^*(t_i),\cdots,s_n^*(t_n))$$

虽然上式有许多的数学符号，但中心思想却非常简单，与完全信息静态博弈中的纳什均衡并没有什么两样：每一个参与者的策略必须是其他参与者策略的最优反应，并且对每一个参与者都如此。因而贝叶斯纳什均衡实际上是贝叶斯静态博弈中的纳什均衡。

现在我们用贝叶斯古诺模型作为例子来解贝叶斯纳什均衡，以加深对上面内容的理解。

已知 $G_B=\{I,A,T,p,u\}$，贝叶斯古诺模型的基本要素有：

(1)参与者：三个垄断企业分别定义为 1、2、3，$I=\{1,2,3\}$。

(2)行动空间：行动 $q_i \in [0,a]$，$A_i=[0, a]$，$i=1,2,3$。

(3)类型空间：$T_i=\{c_L,c_H\}$，$t_{i1}=c_L$，$t_{i2}=c_H$，$i=1,2,3$。

(4)参与者的信念：$p(t_{11})=\theta_1$，$p(t_{12})=1-\theta_1$；$p(t_{21})=\theta_2$，$p(t_{22})=1-\theta_2$；$p(t_{31})=\theta_3$，$p(t_{32})=1-\theta_3$。联合概率分布 $p(t_1,t_2,t_3)$，其中 $p(t_{11},t_{21},t_{31})=\theta_1\theta_2\theta_3$，$p(t_{21},t_{31}|t_{11})=\theta_2\theta_3$，其他的省略。

(5)期望收益函数：参与者 i 的策略 $s_i(t_i)=q_i(t_i)$。由于收益函数完全写出较为繁琐，我们只以参与者 1 的收益函数为例。当参与者 1 类型为低成本，则：

①$p(t_{21},t_{31})\pi_1(q_1,q_2(t_{21}),q_3(t_{31});c_L)$
$=\theta_2\theta_3q_1(a-q_2(c_L)-q_3(c_L)-c_L)$；

②$p(t_{21},t_{32})\pi_1(q_1,q_2(t_{21}),q_3(t_{32});c_L)$
$=\theta_2(1-\theta_3)q_1(a-q_2(c_L)-q_3(c_H)-c_L)$；

③$p(t_{22},t_{31})\pi_1(q_1,q_2(t_{22}),q_3(t_{31});c_L)$
$=(1-\theta_2)\theta_3q_1(a-q_2(c_H)-q_3(c_L)-c_L)$；

①$p(t_{22},t_{32})\pi_1(q_1,q_2(t_{22}),q_3(t_{32});c_L)$
$=(1-\theta_2)(1-\theta_3)q_1(a-q_2(c_H)-q_3(c_H)-c_L)$。

参与者 1 的期望收益函数就是期望利润函数，即：

$$E\pi(q_1,q_2(t_2),q_3(t_3);c_L)=①+②+③+④$$

现在我们求解贝叶斯古诺模型的贝叶斯纳什均衡。根据纳什均衡的定义，应使利润最大化，即：

$$\max_{q_1} E\pi_1(q_1,q_2(t_2),q_3(t_3);c_L)$$

求一阶条件，并整理得：

$$q_1(c_L)=\frac{a-[\theta_2q_2(c_L)+(1-\theta_2)q_2(c_H)]-[\theta_3q_3(c_L)+(1-\theta_3)q_3(c_H)]-c_L}{2} \tag{3.1}$$

利用对称性可以相应得到其他的策略，即：

$$q_1(c_H)=\frac{a-[\theta_2 q_2(c_L)+(1-\theta_2)q_2(c_H)]-[\theta_3 q_3(c_L)+(1-\theta_3)q_3(c_H)]-c_H}{2} \tag{3.2}$$

$$q_2(c_L)=\frac{a-[\theta_1 q_1(c_L)+(1-\theta_1)q_1(c_H)]-[\theta_3 q_3(c_L)+(1-\theta_3)q_3(c_H)]-c_L}{2} \tag{3.3}$$

$$q_2(c_H)=\frac{a-[\theta_1 q_1(c_L)+(1-\theta_1)q_1(c_H)]-[\theta_3 q_3(c_L)+(1-\theta_3)q_3(c_H)]-c_H}{2} \tag{3.4}$$

$$q_3(c_L)=\frac{a-[\theta_1 q_1(c_L)+(1-\theta_1)q_1(c_H)]-[\theta_2 q_2(c_L)+(1-\theta_2)q_2(c_H)]-c_L}{2} \tag{3.5}$$

$$q_3(c_H)=\frac{a-[\theta_1 q_1(c_L)+(1-\theta_1)q_1(c_H)]-[\theta_2 q_2(c_L)+(1-\theta_2)q_2(c_H)]-c_H}{2} \tag{3.6}$$

策略组合$((q_1^*(c_L),q_1^*(c_H)),(q_2^*(c_L),q_2^*(c_H)),(q_3^*(c_L),q_3^*(c_H)))$如果是(3.1)、(3.2)、(3.3)、(3.4)、(3.5)、(3.6)式的解，那么它就是贝叶斯纳什均衡。

需要强调的是，策略是关于行动的一个完全计划，包括了参与者在可能遇到的每一种情况下将选择的可行行动。在贝叶斯古诺模型中，参与者i的策略为$(q_i^*(c_L),q_i^*(c_H))$，即（如果为低成本，选择$q_i^*(c_L)$；如果为高成本，选择$q_i^*(c_H)$）。也许有人认为，要求参与者i的策略包含参与者i每一种可能类型下的可行行动没有必要，毕竟，一旦自然赋予参与者i一特定类型并告知他，参与者就不必再关心如果自然赋予他的是另外一种类型他将如何行动了。例如，自然赋予参与者i低成本，那么参与者i就不用考虑高成本时的情况，也就是策略为$q_i(c_L)$就可以了，而不必要为$(q_i(c_L),q_i(c_H))$而考虑。但问题是，参与者i还需要考虑

另外的参与者将如何行动，其他的参与者对参与者 i 类型并不清楚，只能猜测。而且其他参与者的行动决定于他们对参与者 i 为 T_i 中每一类型 t_i 时，对参与者 i 行动的猜测。所以，即使参与者 i 被自然赋予了某种类型，他仍然要“装作”不知道的样子，去考虑如果他被赋予 T_i 中每一种类型 t_i 时应该如何行动。

更为一般的是，如果一个参与者的策略不包含自然赋予他其他类型时他将如何行动，那么该参与者就不可能求出最优解。因为最优策略是对其他参与者策略的最优反应，在其他参与者并不确切知道该参与者类型的情况下，其他参与者会考虑所有该参与者可能的类型，并以此来决定他们的策略。所以，如果策略不是完全的行动计划，我们就不可能把纳什均衡的概念运用到贝叶斯博弈。在动态博弈中，理解这一点很重要。

混合策略的贝叶斯纳什均衡。

定义 1 在贝叶斯静态博弈 $G_B=\{I,A,T,p,u\}$ 中，参与者 i 的一个混合策略 s_i 是一个由 T_i 映射到 ΔA_i 的函数。ΔA_i 为混合行动空间。$s_i(t_i)$ 代表当参与者 i 的类型为 t_i 时，他所采用的混合策略。对于任意一个行动 $a_i \in A_i$，$s_i(t_i)(a_i)$ 表示在混合策略 $s_i(t_i)$ 下，a_i 被使用的概率。$s=(s_1,s_2,\cdots,s_n)$ 为一个混合策略组合，ΔS_i 为参与者 i 所有可能混合策略的集合，而 $\Delta S=(\Delta S_1,\cdots,\Delta S_n)$ 是所有可能策略组合的集合。

如参与者 i 的类型是 t_i，则参与者 i 在混合策略组合 s 下的收益为：

$$
\begin{aligned}
&\tilde{u}_i(s;t_i) \\
&= \sum_{t_{-i}\in T_{-i}} p(t_{-i}|t_i) \sum_{(a_1,\cdots,a_n)\in A} s_1(t_1)(a_1)\cdots s_n(t_n)(a_n)u_i(a_1,\cdots,a_n;t_{-i},t_i)
\end{aligned}
$$

定义 2 在贝叶斯静态博弈 $G_B=\{I,A,T,p,u\}$ 中，策略组合 $s^*=(s_1^*,\cdots,s_n^*)$ 是一个混合策略贝叶斯纳什均衡，如果对每一个参与者 i 及对 i 的类型空间 T_i 中的每一个类型 t_i，$s_i^*(t_i)$ 满足 $\tilde{u}_i(s_i^*,s_{-i}^*;t_i)\geq \tilde{u}_i(s_i,s_{-i}^*;t_i)$，则对所有的 $s_i\in\Delta S_i$ 成立。

定理 对于有限的贝叶斯静态博弈 $G_B=\{I,A,T,p,u\}$，即参与者

是有限的，A 是有限的，T 是有限的，那么一定存在（混合）贝叶斯纳什均衡。

定理的证明与纳什均衡的证明一样。通过豪尔绍尼转换，我们知道贝叶斯静态博弈可以处理成一串基本式博弈的加权平均——期望值。而每一个基本式博弈都存在着纳什均衡（纳什定理），因而 n 个基本式博弈线性相加（即加权平均）之后的博弈一定存在着纳什均衡。因为线性相加不会改变任何函数的拓扑性质，所以每一个基本式博弈存在纳什均衡，就意味着贝叶斯静态博弈存在纳什均衡，即贝叶斯纳什均衡。实际上，在满足上述定理条件的情况下，在动态博弈中，只要博弈的次数不是无穷的，那么动态博弈中也一定存在纳什均衡。

第三节　拍卖理论

贝叶斯静态博弈最重要的运用就是对拍卖及拍卖机制的分析。最理想的拍卖机制是能够将竞拍人的真实偏好显示出来，而参与者真实偏好的显示在经济学中具有非凡的意义。因为在公共产品配置、外在性解决、企业的委托代理结构、产权配置等重要问题中，核心就在于参与者有动机"虚假"地显示其偏好。也正因为如此，对拍卖及其拍卖机制的分析直接成为信息经济学研究的发端，对拍卖机制作出开创性研究的维克瑞（Vickery）也因此获得了 1996 年诺贝尔经济学奖。

一、拍卖总述

拍卖依投标及付款方式的不同，可分为下列几种形式：

1. 英式拍卖（English auction）

拍卖时拍卖师通常会报出一个起始价格，然后由竞拍人开始报价，后一个报价必须比前一个报价高。当最后一个竞拍人报出的价格高到没有人愿意应拍时，拍卖品就以该价成交。在拍卖所拍卖是英式拍卖最常见的一种拍卖形式，中央电视台黄金时段的广告就是以这种方式竞

拍的(显然这种拍卖近似于完全信息动态博弈,这里我们不妨假设所有参与者同时出价)。

2. 荷式拍卖(Dutch auction)

这种拍卖方式与英式拍卖正相反,因最早在荷兰的拍卖所运用而得名。拍卖师首先会报出一个非常高的价格,如果没有人应拍,竞拍师再报一个较低的价格,如果还没有人应拍,再向下报价,一直到有人应拍为止,拍卖品就以这个价格成交。

3. 密封式一级价格拍卖(Sealed-bid first price auction)

顾名思义,就是所有的竞拍者的报价只有竞拍者本人和拍卖者知道,其他竞拍者不知道,只有报价最高的竞拍者才能中标。[①] 例如,在公共工程招标中,所有竞标者将自己的叫价放在密封的信封中交给发标人,然后发标人公布中标者。

4. 密封式二级价格拍卖(Sealed-bid second price auction)

这和密封式一级价格拍卖的方式一样,也使用保密的方式来竞标,惟一不同的是报价最高的竞拍者中标,但出的价格只是第二高价者的价格。这种拍卖方式是维克瑞发明的,又称维克瑞拍卖。维克瑞 1962 年证明,这种拍卖机制能够诱使竞拍者显示他的真实偏好,作为诱使竞拍者显示真实偏好的代价就是拍卖者只能得到第二高价,第一高价与第二高价的差价就是对竞拍者说真话的奖励。这实际上意味着,说真话和帕累托最优不可能同时达到,换句话说,就是激励兼容问题。激励兼容问题是整个信息经济学的基石。

拍卖根据投标者拥有的信息不同,又可分为下列几种情况:

1. 私人价值拍卖(Private value auction)

这是指竞拍者对拍卖品价值的评估是自己的私人信息,别人不知道,除非他说出来。在现实生活中,物品的价值评估实际上都是私人信息,市场机制的作用就在于通过需求和供给机制形成市场价格,所以市场机制又被理解为拍卖人机制,通过市场这个特殊的拍卖场所发现(公共)价格。

---

① 在公共工程招标中,是最低报价的投标人中标,这时所有投标人的报价可以取负值。

2. 公共价值拍卖(Common value auction)

在有些特定的情况下，拍卖品的价值成了公共信息，这时拍卖就成了所谓的公共价值拍卖。例如，有一处房产的市场价值为 20 万(因为与它同质的其他房产市场上卖 20 万)，但它的主人破产需要拍卖房产抵债，这时房产拍卖就是公共价值拍卖，因为拍卖价格绝对不会超过 20 万，也不会低于 20 万。

3. 相关价值拍卖(Correlated value auction)

当竞拍者之间对拍卖品的价值评估具有相关性时，拍卖就是相关价值拍卖。例如，当某人对拍卖品的价值评价越高，他预期别人对此物的价值评价也越高。竞标者之间秘密勾结也会形成相关价值拍卖。

上述三种信息结构的拍卖，我们主要分析第一种。

一个有趣的问题是，四种拍卖机制在均衡上各有什么特点，哪一种拍卖机制对拍卖者更有利。出人意料的是，在私人价值的假设下四种拍卖制度完全等价，这就是有名的等报酬定理(Revenue equivalent theorem)。

定理 3.1　等报酬定理　　在私人价值拍卖里，假设拍卖品对第 i 个竞拍者的价值为 v_i，$i=1,2,\cdots,n$，且每一个 v_i 的值，都是从$[0,\bar{v}]$中用 $f(v)$这个概率密度函数随机独立抽取的。假设$f(v)>0$，$\forall\ v\in[0,\bar{v}]$，则拍卖者或每一个竞拍者的预期报酬，在上述四种拍卖制度下都是一样的。

我们将用下面三个命题来证明这个定理。为了简单一些，假设竞标者只有两人，实际上由于 $f(v)$是独立的，所以两人的分析和 n 人的分析完全一样，这时把 $f(v)$理解为联合密度函数即可。

命题 3.1　荷式拍卖和密封式一级价格拍卖是完全相同的。

证明：上述问题等价于证明任何一个竞拍者，他在荷式拍卖和密封式一级价格拍卖下所面对的情况完全一样。

一个竞拍者在荷式拍卖会场中的策略是，在尚未有人得标之前，他应该在价格下降到何处时标购。而密封式一级价格拍卖中的竞拍者所要考虑的是他应该出多高的价格。这两种策略的形式实际上是一样的。也就是说，在这两种制度之下，竞标者的策略都是在给定他自己的价值

v 时，他愿意出的价格。令 $b_i(v_i)$ 是竞拍者 $i(i=1,2)$，当他认为标的物的价值为 v_i 时，所愿意竞拍的价格。这可以很容易证明，在均衡时，每一个人的出价函数 $b_i(v_i)$ 一定是 v 的增函数。假设出价函数是对称均衡的，即所有竞标者的出价函数都一样，$b_i(v)=b(v)$ 对任意 i 成立。

由于这种物品对 i 的个人价值为 v_i，因此他的标价为 $b(v_i)$，又由于函数 $b(v)$ 是 v 的增函数，因此若 i 的对手对该物品的个人价值小于 $b^{-1}(b(v_i))$，则 i 可以得标，获标的概率为（当 $v_i=v_j$ 时，两人抛硬币决定标的物归谁）：

$$F(b(v)<b(v_i)\,|\,b(v_i))=F(v<v_i\,|\,v_i)=\int_0^{b^{-1}(b(v_i))}f(v)\mathrm{d}v$$

在 $b(v_i)$ 的策略下，如果 i 得标，那么他的报酬为 $v_i-b(v_i)$。因此 i 的期望报酬为：

$$\int_0^{b^{-1}(b(v_i))}(v_i-b(v_i))f(v)\mathrm{d}v$$

在荷式拍卖里，若价格降至 y 时标的物仍未出售，表示 i 的对手对这件物品的价值评估低于 $b^{-1}(y)$。因此投标者 i 会利用贝叶斯法则，来重新修正对手价值的概率密度函数。由于 v 低于 $b^{-1}(y)$ 的可能性为 $F(b^{-1}(y))$，所以用贝叶斯法则修订之后的新密度函数为 $f(v)/F(b^{-1}(y))$。因此，如果 i 的策略是 $b(v_i)$，且在价格降到 y 时还没有标走，那么他用 $b(v_i)$ 这个策略的预期报酬是：

$$\int_0^{b^{-1}(b(v_i))}(v_i-b(v_i))\frac{f(v)}{F(b^{-1}(y))}\mathrm{d}v$$

由于 $1/F(b_i^{-1}(y))$ 是固定常数，所以竞拍者在荷式拍卖与密封式一级价格拍卖里所面对的目标函数，只差一个常数。因此，对竞标者来说两者完全一样。

推论 3.1 在英式拍卖里，$b(v)=v$ 是一个优策略。

证明：竞标者不可能出价 $b(v)>v$，因为这等于是赔本买卖。如竞标者的策略是 $b(v)<v$，那么有两种可能情况：

(1)如果出价 $b(v)=v$ 得标，那么让 $b(v)<v$ 时，可能会得标，也可能不会得标。如果得标，那么他所付出的价钱 $b(v)=v\#$ 和 $b(v)=v$ 时的情景完全相同，都是付出他之前最后一人放弃时的价格。因此 $b(v)=v$ 和 $b(v)=v\#$ 一样好。如他出价 $b(v)=v\#$ 不能得标，那么这个策略的效用会比 $b(v)=v$ 的策略要差。因此 $b(v)=v$ 的策略比 $b(v)=v\#$ 要好。

(2)如果 $b(v)=v$ 本来就不会得标，那么 $b(v)=v\#$ 也不可能得标，因此两个策略是一样的。综上所述，可知 $b(v)=v$ 相对于所有 $b(v)\neq v$ 是一个优策略。

推论 3.2　在密封式二级价格拍卖里，如标的物对某一个竞标者的价值为 v，则其均衡投标价格为 $b(v)=v$。

证明：假设竞标者的策略为 $b(v)$，那么他的预期报酬为：

$$\int_0^v (v-b(v'))f(v')\mathrm{d}v' = vF(v) - \int_0^v b(v')f(v')\mathrm{d}v' \quad (3.1)$$

其中 $b(v')$ 是标的物价值对其为 v' 的人投标的价格(即第二高价)。如果 $b(v)$ 是最优策略，那么 b 在任何一个 v 下都是最优的，因此 b 在每一点都要满足一阶微分条件。将(3.1)式对 v 微分，可得 $vf(v)-b(v)f(v)=0$。[①] 因此 $b(v)=v$。

命题 3.2　英式拍卖和密封式二级价格拍卖都产生相同的预期报酬。

证明：由推论 3.1 和推论 3.2 可知。

命题 3.3　不论是卖方还是买方，密封式一级价格和二级价格拍卖者的预期报酬是一样的。

证明：在二级价格拍卖里，假设标的物对两个竞标者的价值是 v 和 v'。如果 $v>v'$，那么卖方得到的价钱是第二高价的 $b(v')$，而竞标者会出价 v 和 v' 的概率密度函数各是 $f(v)$ 和 $f(v')$。由于每一个人的出价

---

① 在对(3.1)式的第一项 $vF(v)$ 微分时，我们只对 $F(v)$ 这项微分，而不对 v 这项微分，因为 v 是竞标者的真实价值，是一个常数。换言之，不论策略如何改变，标的物对他的真实价值是不变的 v。

都是独立的，所以有人同时出价 v 和 v' 的概率密度将是 $f(v)$ 和 $f(v')$ 的乘积。将所有 0 到 v 可能出价 $b(v')$ 的收益积分起来，便是卖方的预期收益。因此，卖方预期收益为：

$$\int_0^{\bar{v}}\int_0^{v} b(v')f(v')f(v)\mathrm{d}v'\mathrm{d}v = \int_0^{\bar{v}} f(v)\int_0^{v} b(v')f(v')\mathrm{d}v'\mathrm{d}v$$
$$= \int_0^{\bar{v}} f(v)\int_0^{v} f(v')v'\mathrm{d}v'\mathrm{d}v \quad (3.2)$$

上式中，第二个等式的来源是在二级价格拍卖里，均衡出价 $b(v)=v$（推论 3.2）。至于买方的预期收益，假设他对于标的物的认定价值是 v，那么他的出价为 $b(v)$。由于他只需要付第二高的价格，而这个出价 $b(v')$ 出现的概率密度函数为 $f(v')$，且此时的收益为 $(v-b(v'))$，因此预期收益为：

$$\int_0^{v}(v - b(v'))f(v')\mathrm{d}v' = \int_0^{v}(v - v')f(v')\mathrm{d}v'$$
$$= vF(v) - \int_0^{v} v'f(v')\mathrm{d}v' \qquad (3.3)$$

接下来，我们求第一高价拍卖下的均衡叫价 $b(v)$。假设某一个竞标者的认定价值为 v，由于他叫价 $b(v)$，因此他的预期收益是：

$$\int_0^{v}(v - b(v))f(v')\mathrm{d}v' = (v - b(v))F(v)$$

在最优策略下，$b(v)$ 必须在每一个 v 下满足一阶条件，因此 $(v-b(v))f(v)-b'(v)F(v)=0$，得 $vf(v)=b(v)f(v)+b'(v)F(v)=\mathrm{d}(b(v)F(v))/\mathrm{d}v$，对两边积分可得到 $\int_0^{v} v'f(v')\mathrm{d}v' = b(v)F(v)$。因此 $b(v) = (\int_0^{v} v'f(v')\mathrm{d}v')/F(v)$。由此我们可计算出卖方在一级价格拍卖的预期收益为：

$$\int_0^{\bar{v}}\int_0^{v} b(v)f(v)f(v')\mathrm{d}v' = \int_0^{\bar{v}} b(v)f(v)F(v)\mathrm{d}v$$

$$= \int_0^v \left(\int_0^v v' f(v') \mathrm{d}v' / F(v) \right) f(v) F(v) \mathrm{d}v$$

$$= \int_0^v f(v) \int_0^v v' f(v') \mathrm{d}v' \mathrm{d}v \tag{3.4}$$

将(3.2)和(3.4)比较，我们马上知道一级价格和二级价格拍卖的卖方预期利润值是一样的。

买方的预期利润为：

$$\begin{aligned} \int_0^v (v - b(v)) f(v') \mathrm{d}v' &= (v - b(v)) F(v) \\ &= vF(v) - b(v)F(v) \\ &= vF(v) - \frac{\int_0^v v' f(v') \mathrm{d}v'}{F(v)} F(v) \\ &= vF(v) - \int_0^v v' f(v') \mathrm{d}v' \end{aligned} \tag{3.5}$$

比较(3.3)和(3.5)可知一级价格和二级价格拍卖对买方的预期利润也一样。

综上所述，命题得证。

由命题 3.3 我们可知，一级价格和二级价格拍卖，卖方的预期收益和买方的预期收益是一样的。但由于一级价格拍卖和荷式拍卖完全相同(命题 3.1)，而二级价格拍卖和英式拍卖买卖双方的预期收益一样(命题 3.2)，因此我们证明了这四种拍卖方式对卖方、买方的预期收益是相同的。

等报酬定理只有在私人价值拍卖下，每一个人的价值是独立抽取时才成立。在更一般的情况下，英式拍卖预期收益最高，密封式一级价格拍卖预期收益最低。米尔格罗姆(Milgrom)和韦伯(Weber)在 1982 年证明了这个结果。①

定理 3.2　拍卖机制有下列关系：

---

① Milgrom, P. and Weber, R. J. 1982: "A Theory of Auctions and Competitive Bidding", *Economitrica* 50, pp. 1089～1122.

(1)荷式拍卖和密封式一级价格拍卖完全等价;

(2)在英式拍卖中,卖方的预期报酬大于密封式二级价格拍卖;

(3)在密封式二级价格拍卖中,卖方的预期报酬大于密封式一级价格拍卖。

由定理3.2我们可以知道,在四种拍卖机制中,英式拍卖是预期收益最多的一种拍卖机制。这就是为什么在一般拍卖所里,最常见到英式拍卖。既然如此,为什么密封式一级价格拍卖,也是常见的一种拍卖机制呢?采用密封招标的用意,其实不在预期收益的多寡,而在于防止竞标者之间的勾结。在拍卖的过程中,常会发生竞标者相互勾结的情形。这以政府经常性的招标最常见。例如,在工程招标中,营造公司可以相互决定投标价格,然后以非常高的造价让某一家厂商去包下一个工程。又如,在政府的财产拍卖时,投标者也可以相互约定让某一投标者以极低的价格中标。这就是投标勾结。但由于投标勾结者之间的约定,没有法律上的效力,也没有很强的约束力,因此这种勾结要成功必须借助拍卖制度本身的缺陷。如果投标的过程使用英式拍卖,那么任何一家厂商想要欺骗其他厂商非常困难。比如说,如果他们之间约好用P的价格让某一家厂商得标,而如果这个价格高于成本很多,那么由于利润很高,其他厂商就有违背约定的动机。但是在英式拍卖里,其他厂商能当场发现该厂商违反事前约定,而用更低的价格来反制,使得违反约定的图谋难以得逞。也就是说,在英式拍卖机制下,任何一个勾结厂商想违反勾结约定去欺骗其他厂商较为困难。正因为如此,这种制度反而增强了厂商相互勾结的可能性。但在密封制度下,由于厂商无法知道其他厂商的标价,因此勾结较难实现。为了有效防止竞标者的勾结,密封式制度是较好的一种制度安排。从竞拍者本身来看,勾结本身也需要一些条件。较有可能产生勾结的拍卖市场,是竞标者数目较少的场所,且经常举行拍卖活动。因为达成和维持一个勾结协定的难度随竞标者数目的增加而成几何数增加。这就是所谓的十个监督者不如一个竞争者。在工程招标中,一是由于营造厂商的数目较少,二是由于工程招标活动较为频繁,因此勾结的可能性很大。这就是为什么公共工程招标多采用密封式拍卖制度的原因。

二、拍卖例子

为了更好地理解拍卖理论，我们具体举一些例子。

1. 密封式一级价格拍卖

假设有两人竞标，中标方式为密封式一级价格拍卖，即他们把各自的投标价装入信封，价高者中标。投标人 i 对商品的估价为 v_i，投标人 i 的叫价为 b，如中标则 i 的收益为 v_i-b。两个投标人对标的物的估价相互独立，并服从[0,1]区间上的均匀分布。投标价格不能为负，且双方同时给出各自的投标价。出价较高的一方得到商品，并支付他报的价格；另一方的收益和支付都为 0。在投标价相等的情况下，由掷硬币决定谁中标。假设投标者是风险中性的，所有以上信息都是共同信息。

要把该博弈转化为贝氏博弈的基本式，必须确定行动空间、类型空间、信念及收益函数。定义该博弈的基本式为 $G_B=\{I,A,T,p,u\}$，其中：

(1)参与者：两个投标人(Bidder)分别定义为 1,2，$I=\{1,2\}$。

(2)行动空间：参与者 i 的行动是给出一个非负的投标价 b_i，行动空间 $A_i=[0,+\infty)$，$i=1,2$。

(3)类型空间：投标人的类型即他的估价 v_i，类型空间 $T_i=[0,1]$，$i=1,2$。

(4)信念：已知密度函数 $f(v_i)=1$，由于 v_i 是相互独立的，所以根据贝叶斯法则，参与者 i 推断 v_j 服从[0,1]区间上的均匀分布，而不依赖于 v_i 的值，即 $f(v_j|v_i)=f(v_j)$。

(5)收益函数：参与者 i 的收益函数为：

$$u_i(b_1,b_2;v_1,v_2)=\begin{cases} v_i-b_i & b_i>b_j \\ (v_i-b_i)/2 & b_i=b_j \\ 0 & b_i<b_j \end{cases}$$

为推导这一博弈的贝叶斯纳什均衡，必须用行动空间和类型空间构建参与者的策略空间。前面已讲过，在静态贝叶斯博弈中，一个策略是从类型映射到行动的函数。从而，参与者 i 的一个策略为函数 $b_i(v_i)$，据此可以决定 i 在每一种类型(即对商品的估价)下选择的投标价格。

如果策略组合$(b_1(v_1),b_2(v_2))$是贝叶斯纳什均衡，那么对[0,1]中的每一个v_i，$b_i(v_i)$一定满足：

$$\max_{b_i}(v_i-b_i)\mathrm{Prob}\{b_i>b_j(v_j)\}+\frac{1}{2}(v_i-b_i)\mathrm{Prob}\{b_i=b_j(v_j)\}$$

要解出贝叶斯纳什均衡必须明确给出出价函数$b_i(v_i)$。这里我们不妨考虑最简单的一种出价策略，即出价函数为线性函数。实际上出价函数本身也需要从整个博弈来考虑是不是一种最优的制定策略的方法，因为$b_i(v_i)$可能为二次函数、三次函数，等等，而不一定是线性函数。但为了使问题简单，我们不妨先假设该博弈存在一组线性均衡解$b_1(v_1)=a_1+c_1v_1$及$b_2(v_2)=a_2+c_2v_2$。通过分析，我们会发现由于参与者的估价是均匀分布的，这样的线性均衡解不仅存在，而且相对于其他出价函数的均衡要优，其结果为$b_i(v_i)=v_i/2$。下面我们就证明这一点。

假设参与者j采取策略$b_j(v_j)=a_j+c_jv_j$，对于一个给定的v_i值，参与者i的最优反应为下式的解：

$$\max_{b_i}(v_i-b_i)\mathrm{Prob}\{b_i>a_j+c_jv_j\}$$

因为$\mathrm{Prob}\{b_i=b_j(v_j)\}=0$，[①]所以我们省略掉了第二项。显然参与者$i$的投标价不可能低于参与者$j$最低的可能投标价格$a_j$，也不可能高于$j$最高的可能投标价格，所以有$a_j\leqslant b_i\leqslant a_j+c_j$，于是：

$$\mathrm{Prob}\{b_i>a_j+c_jv_j\}=\mathrm{Prob}\{v_j<\frac{b_i-a_j}{c_j}\}=\frac{b_i-a_j}{c_j}$$

因而参与者i的最优反应策略为：

$$b_i(v_i)=\begin{cases}(v_i+a_j)/2 & v_i\geqslant a_j\\ a_j & v_i<a_j\end{cases}$$

上述思想非常简单，只有$v_i>a_j$，投标人i才会去投标；如果$v_i<a_j$，那

① 因为b是连续性随机变量，所以$\mathrm{Prob}\{b_i=b_j\}=0$。$b=a+cv$，$v$服从均匀分布，$b$也服从均匀分布。

么 i 的叫价 b_i 实际上小于等于 a_j，即 i 不可能中标。这里我们不妨取等号。

如果 $0<a_j<1$，则一定存在某些 v_i 的值，使得 $v_i<a_j$，这时 $b_i(v_i)$就不可能是线性的。由于我们要寻找线性的均衡，这就可以排除$0<a_j<1$的情况，而只讨论 $a_j\geqslant1$ 及 $a_j\leqslant0$。但前一种情况是不可能在均衡中出现的，因为估价较高一方对投标价的最优选择是不低于估价较低一方的投标价，我们有 $c_j\geqslant0$，但这时 $a_j\geqslant1$ 便意味着 $b_j(v_j)\geqslant v_j$，即出价超过估价，而这肯定不会是最优的。因此，如果要求 $b_i(v_i)$是线性的，则一定有 $a_j\leqslant0$，这种情况下 $b_i(v_i)=(v_i+a_j)/2$，于是可得 $a_i=a_j/2$ 及 $c_i=1/2$。

我们可以假定参与者 i 采取策略 $b_i(v_i)=a_i+c_iv_i$，对参与者 j 重复上面的分析，得到类似的结果 $a_i\leqslant0$，$a_j=a_i/2$ 以及 $c_j=1/2$。解这两组结果构成的方程组，可得 $a_i=a_j=0$ 和 $c_i=c_j=1/2$，$b_i(v_i)=v_i/2$。这就证明了我们前面的结论。

一个自然的想法是，在密封式一级价格博弈中是否还存在另外的贝叶斯纳什均衡？如果投标者估价的概率分布发生变化，均衡的投标价格将如何变化？运用前面的方法（即先假定一个线性策略，再推导出使策略符合均衡条件的系数）已不能回答这两个问题，因为其他均衡中 b_i 的函数形式具有无穷多种，试图猜测这一博弈其他均衡中 b_i 的函数形式是徒劳的，并且当估价服从任何其他分布时，线性均衡均不存在。不过幸运的是，当估价 v_i 服从均匀分布时，存在线性均衡。

下面我们考虑另外一种情况—— 出价函数是对称均衡的，即所有投标者的出价函数都一样，$b_i(v)=b(v)$对任意 i 成立，但估价 v_i 仍服从均匀分布。在参与者策略严格递增及可微的假定下，可证明惟一的对称贝叶斯纳什均衡就是本节推导出的线性均衡。这里的分析很容易扩展到较广类型的估价分布，以及两个以上投标者的情况。

假设参与者 j 采取策略 $b(\cdot)$，同时假定 $b(\cdot)$严格递增并可微，则对一个给定的值 v_i，参与者 i 的最优投标价格应满足：

$$\max_{b_i}(v_i-b_i)\text{Prob}\{b_i>b(v_j)\}$$

令 $b^{-1}(b_j)$表示参与者 j 在选择投标价格 b_j 时所持有的估价，即如

果 $b_j=b(v_j)$，则 $b^{-1}(b_j)=v_j$。由于 v_j 服从区间[0,1]上的均匀分布，$b_i>b(v_j)$ 的概率等于 $b^{-1}(b_i)>v_j$ 的概率，后者又等于 $b^{-1}(b_i)$，因为 $\text{Prob}\{b_i>b(v_j)\}=\text{Prob}\{b^{-1}(b_i)>v_j\}=b^{-1}(b_i)$。因而参与者 i 最优化问题的一阶条件为：

$$-b^{-1}(b_i)+(v_i-b_i)\frac{\mathrm{d}}{\mathrm{d}b_i}b^{-1}(b_i)=0$$

上面的一阶条件在给定投标者 i 的估价 v_i 时，是关于投标者 i 对投标者 j 的策略 $b(\cdot)$最优反应的隐函数。如果要使 $b(\cdot)$成为对称的贝叶斯纳什均衡，我们要求一阶条件的解应该等于 $b(v_i)$。也就是说，只要投标方 j 也选择同一策略，对参与者 i 每一可能的估价 v_i，投标人 i 都不希望背离策略 $b(\cdot)$。为加入这一要求，我们用 $b_i=b(v_i)$代入一阶条件，得：

$$-b^{-1}(b(v_i))+(v_i-b(v_i))\frac{\mathrm{d}}{\mathrm{d}b_i}b^{-1}(b(v_i))=0$$

上式 $b^{-1}(b(v_i))$就是 v_i，而且 $\frac{\mathrm{d}\{b^{-1}(b(v_i))\}}{\mathrm{d}b_i}=\frac{1}{b'(v_i)}$。$\frac{\mathrm{d}\{b^{-1}(b_i)\}}{\mathrm{d}b_i}$衡量的是如果投标价格发生变动时，投标者 i 的估价 v_i 可能发生的变化；而 $1/b'(v_i)$衡量的是如果估价发生单位变化，其投标价格将随之发生多大变动。因此，$b(\cdot)$必须满足一阶微分方程，即：

$$-v_i+(v_i-b(v_i))\frac{1}{b'(v_i)}=0$$

可以将其简化表示为 $b'(v_i)v_i+b(v_i)=v_i$，微分方程等式的左边恰好等于 $d\{b(v_i)v_i\}/dv_i$，对左右两方同时求不定积分得：

$$\int[b'(v_i)v_i+b(v_i)]\mathrm{d}v_i=\int v_i\mathrm{d}v_i$$

$$b(v_i)v_i=\frac{1}{2}v_i^2+k$$

其中 k 为常数。要消除 k，我们需要一个边界条件。根据已知条件，没有参与者愿意出高于自己估价的投标价格。因而，我们要求对所有的

v_i，都有 $b(v_i)<v_i$。其中一个特例是当 $v_i=0$ 时，我们要求 $b(0)\leqslant 0$。由于投标价格被限制为非负，这意味着 $b(0)=0$，于是 $k=0$，因而 $b(v_i)=v_i/2$，即无论出价函数是否对称，出价都为估价的一半。

根据常识而言，投标人越多，投标人的出价和估价的差异应当越小。假设有 n 个投标人，其他假设同上，那么在给定第 i 个参与者的策略 $b(\cdot)$，对一个给定的值 v_i，参与者 i 的最优投标价格应满足：

$$\max_{b_i}[v_i - b(v_i)] \times \prod_{j\neq i}\text{Prob}\{b(v_i) > b(v_j)\}$$

或　$$\max_{b_i}(v_i-b_i)[F(b_i)]^{n-1}$$

其中 $F(b_i)=\text{Prob}\{b_i>b_j\}$。

根据对称性和前面的计算可知 $F(b_i)=\text{Prob}\{b_i>b_j\}=v_i$，一阶条件为：

$$-[F(b_i)]^{n-1}+(v_i-b_i)(n-1)[F(b_i)]^{n-2}\times\frac{\mathrm{d}}{\mathrm{d}b_i}F(b_i)=0$$

$$-F(b_i)+[v_i-b(v_i)](n-1)F'(b_i)=0$$

$$-F(b_i)+[F(b_i)-b_i](n-1)F'(b_i)=0$$

解上述一阶微分方程得：

$$b(v_i)=\frac{n-1}{n}v_i$$

这说明，当 n 趋于无穷大时，投标者 i 的出价等于估价。

2. 双向拍卖

下面我们分析双向拍卖的交易博弈，买方和卖方对自己的估价都是私人信息。买方确定一个买价 p_b，同时卖方给出一个卖价 p_s。如果 $p_b\geqslant p_s$，则交易以 $p=(p_b+p_s)/2$ 的价格进行；如果 $p_b<p_s$，则不发生交易。买方对标的物的估价为 v_b，卖方的估价为 v_s，双方的估价都是私人信息，并且服从[0,1]区间的均匀分布。如果买方以价格 p 购得商品，则可获得 v_b-p 的收益，卖方获得 $p-v_s$ 的收益；如果交易未成功，买卖双方获得的效用为零。

要解出这一贝叶斯静态博弈的贝氏均衡，首先要进行规范化处理，

令 $G_B=\{I,A,T,p,u\}$，其中的五个要素请读者自己给出。其次，根据行动空间和类型空间定义买卖双方的策略，买方的策略定义为函数 $p_b(v_b)$，即在每一个可能的类型下将会给出的买价。类似地，卖方的策略为 $p_s(v_s)$。

根据贝叶斯纳什均衡（以后简称贝氏均衡）的定义，如果 $(p_b(v_b),p_s(v_s))$ 为贝氏均衡，那么在给定卖者的出价期望值下，对 $[0,1]$ 区间内的每一个类型 v_b，$p_b(v_b)$ 应满足：

$$\max_{p_b}\{v_b-\frac{p_b+E[p_s(v_s)\mid p_b\geqslant p_s(v_s)]}{2}\}\text{Prob}\{p_b\geqslant p_s(v_s)\} \quad (3.6)$$

其中 $E[p_s(v_s)\mid p_b\geqslant p_s(v_s)]$ 为在卖方价格小于买方价格的条件下，卖方价格的期望值。

在给定买者的出价期望值下，对 $[0,1]$ 区间内的每一个 v_s，$p_s(v_s)$ 应满足：

$$\max_{p_s}\{\frac{p_s+E[p_b(v_b)\mid p_b(v_b)\geqslant p_s]}{2}-v_s\}\text{Prob}\{p_b(v_b)\geqslant p_s\} \quad (3.7)$$

其中 $E[p_b(v_b)\mid p_b(v_b)\geqslant p_s]$ 为在买方价格大于卖方价格的条件下，买方价格的期望值。

正像我们前面讲到的，出价函数 p 具有无穷多个，因而该博弈的贝氏均衡可能具有无穷多个。这里我们考虑两个特殊的均衡：一个是单一价格均衡，一个是线性均衡。由于类型 v 服从均匀分布，所以一定存在线性均衡。

所谓单一价格均衡，就是要么以单一的价格成交，要么不成交。设 x 是属于区间 $[0,1]$ 上的任意一个值，令买方的策略为：如果 $v_b\geqslant x$，则出买价 x，其他情况下出价为 0；同时令卖方的策略为：如果 $v_s\leqslant x$，则出卖价为 x，其他情况下出卖价为 1。注意，在其他情况下，买方出价为 0，卖方出价为 1（v 是 $[0,1]$ 区间上的均匀分布），所以不可能成交。

给定买方的策略，卖方只能在以价格 x 成交或不成交之间进行选择，在 $v_s\leqslant x$ 下成交，是卖方的严格优策略，而在 $v_s\geqslant x$ 下不成交，同样是卖方的严格优策略。所以在 $v_s\leqslant x$ 时出卖价为 x，其他情况下出卖价

为 1，是卖方对买方策略的最优反应。反之，在给定卖方策略的情况下，买方也只能在价格 x 成交或不成交之间选择，在 $v_b \geq x$ 下成交，是买方的严格优策略，而在 $v_b \leq x$ 下不成交，是买方的严格优策略。所以在 $v_b \geq x$时出买价 x，其他情况下出价 0，是买方对卖方策略的最优反应。因而单一价格策略为该博弈的一个贝氏均衡。在这一均衡结果下，图 3.5 标出了所有可能发生交易的区域。从微观经济学的知识来讲，如果 $v_b > v_s$，那么交易都应当发生，因为这时交易能够改善双方的福利（例如以 $(v_b+v_s)/2$ 的价格成交），因而交易是有效率的。但图中斜线阴影部分虽满足效率条件，却没有发生交易。

现在我们推导双向拍卖的线性均衡，仍如上节强调的，我们并没有限制策略空间只包含线性均衡，而是允许参与者任意选择策略，看是否存在一个线性均衡。由于估价服从均匀分布，线性均衡是存在的。除了单一价格均衡和线性均衡之外，博弈还存在着许多其他的均衡，但线性均衡有着很好的效率特征。

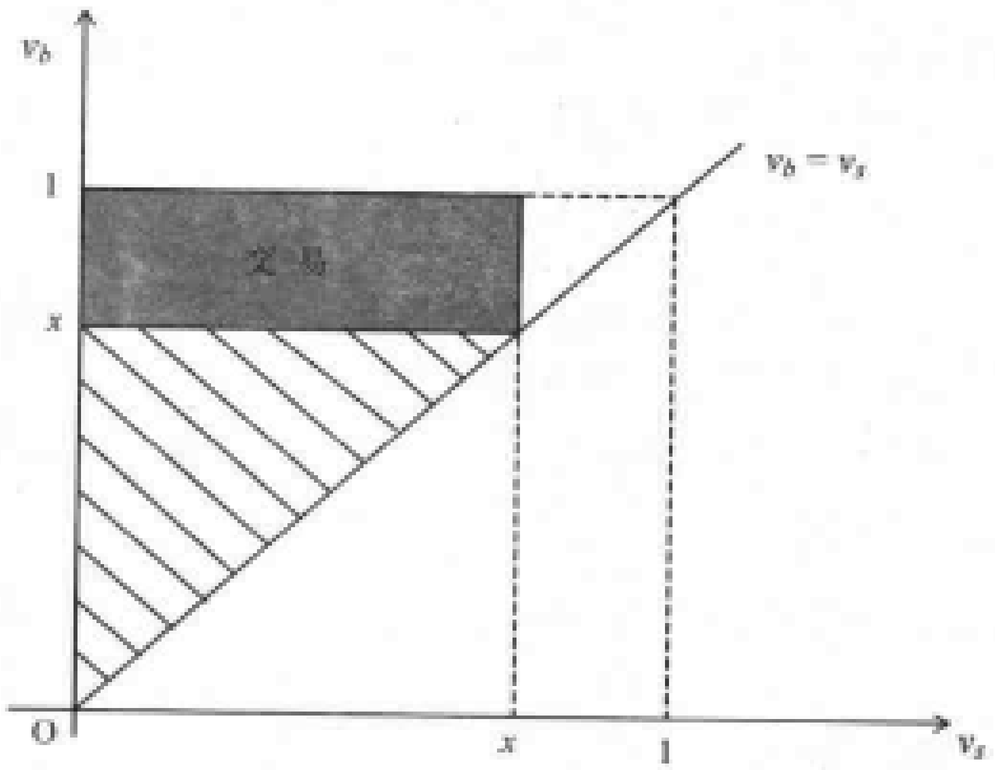


图 3.5　单一价格均衡

假设卖方的策略为 $p_s(v_s)=a_s+c_sv_s$，则 p_s 服从区间 $[a_s,a_s+c_s]$ 上的均匀分布，于是(3.6)式可化为：

$$\max_{p_b}\left[v_b-\frac{1}{2}\left(p_b+\frac{a_s+p_b}{2}\right)\right]\frac{p_b-a_s}{c_s}$$

由于 p_s 服从均匀分布，在 $p_b\geqslant p_s(v_s)$ 的条件下，p_s 的最小值为 a_s，最大值为 p_b，所以 $E[p_s(v_s)\mid p_b\geqslant p_s(v_s)]=(a_s+p_b)/2$。

由上式的一阶条件可推出：

$$p_b=\frac{2}{3}v_b+\frac{1}{3}a_s \tag{3.8}$$

相似地，假设买方的策略为 $p_b(v_b)=a_b+c_bv_b$，则 p_b 服从区间 $[a_b,a_b+c_b]$ 上的均匀分布，于是(3.7)式可化为：

$$\max_{p_s}\left[\frac{1}{2}\left(p_s+\frac{p_s+a_b+c_b}{2}\right)-v_s\right]\frac{a_b+c_b-p_s}{c_b}$$

由于 p_b 服从均匀分布，在 $p_b(v_b)\geqslant p_s$ 的条件下，p_b 的最小值为 p_s，最大值为 a_b+c_b，所以 $E[p_b(v_b)\mid p_b(v_b)\geqslant p_s]=(p_s+a_b+c_b)/2$。

由上式的一阶条件可得：

$$p_s=\frac{2}{3}v_s+\frac{1}{3}(a_b+c_b) \tag{3.9}$$

这表明，如果卖方选择一个线性策略，则买方最优的策略也是线性的，反之亦然，即存在一个线性均衡。

由(3.8)式可知 $a_b+c_bv_b=p_b=\frac{2}{3}v_b+\frac{1}{3}a_s$，从而得 $c_b=2/3$，$a_b=a_s/3$。

由(3.9)式可知 $a_s+c_sv_s=p_s=\frac{2}{3}v_s+\frac{1}{3}(a_b+c_b)$，从而得 $c_s=2/3$，$a_s=(a_b+c_b)/3$。

那么，线性均衡策略为：

$$p_b(v_b)=\frac{2}{3}v_b+\frac{1}{12} \tag{3.10}$$

和　$p_s(v_s)=\frac{2}{3}v_s+\frac{1}{4}$　(3.11)

[(3.10),(3.11)]即为双向拍卖的一个贝叶斯纳什均衡。

前面讲过，双向拍卖中交易发生的必要条件是$p_b \geqslant p_s$。如图3.6所示，当卖方的估价高于3/4时，他的要价超过了买方的最高可能出价$p_b(1)=3/4$；当买方的估价低于1/4时，他的出价低于卖方的最低可能要价$p_s(0)=1/4$。在上述两种情况下，交易都不会发生。这意味着，在线性均衡中，只有当$v_b \geqslant v_s+(1/4)$时才会有交易发生，如图3.7所示。比较图3.5和图3.7，可清楚地看出在单一价格均衡及线性均衡中交易发生所要求的估价组合。在这两种情况下，交易的潜在价值最大时(具体讲，当$v_s=0$，$v_b=1$)都会发生交易。但是，单一价格均衡漏掉了一些有价值的交易，比如$v_s=0$，$v_b=x-\varepsilon$，其中ε是足够小的正数，而包含了一些几乎没有什么价值的交易，比如$v_s=x-\varepsilon$且$v_b=x+\varepsilon$。相反，在线性均衡中，过滤掉了所有价值不大的交易，只包含了价值至少在1/4以上的交易。这表明从参与者可得到的期望收益的角度来看，线性均衡要优于单一价格均衡，但是否最优还需要考察其他的均衡。幸运的是线性均衡确实是最优的。

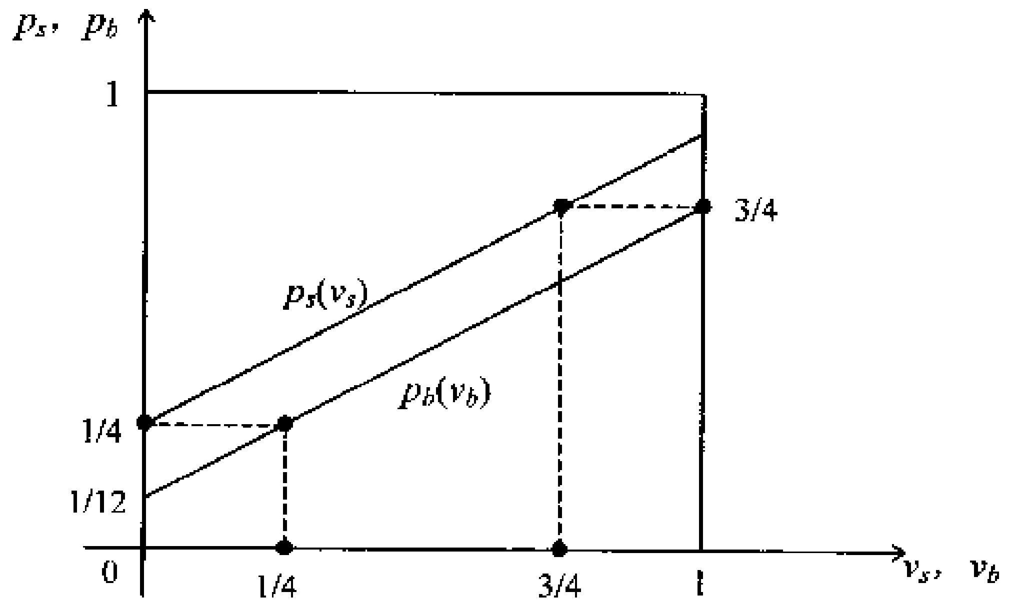


图3.6　线性均衡下的可能交易

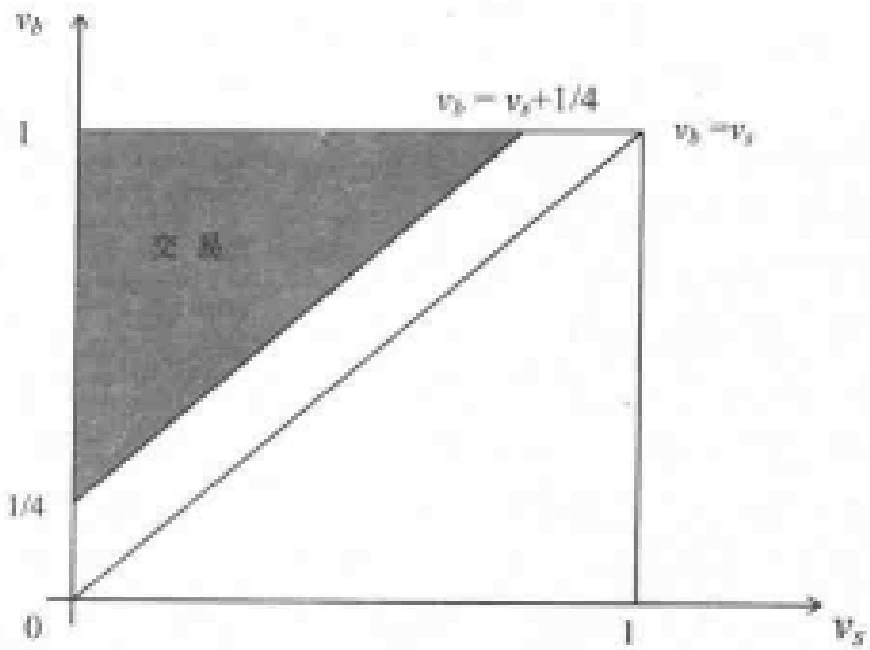


图 3.7 线性均衡下的交易区间

前面我们讲过，在类型服从均匀分布的情况下，线性均衡不仅存在，而且相对于其他均衡是最优的。迈尔森（Myerson）和萨特思韦特（Satterthwaite）在 1983 年证明了这一结论：对于这里考虑的估价的均匀分布，在双向拍卖的线性均衡中，参与者的期望收益要高于其他任何贝氏均衡。这意味着在双向拍卖中，不存在这样的贝氏均衡：交易当且仅当有效率时（$v_b \geq v_s$）必定发生（因为线性均衡过滤掉了一些价值不大的交易）。他们进一步证明了这一结果的普遍性，即：

定理 3.3 无效率定理　　如果 v_b 在区间$[x_b, y_b]$连续分布，v_s 在区间$[x_s, y_s]$连续分布，其中 $y_s > x_b$ 且 $y_b > x_s$，那么在讨价还价博弈中，仅当有效率时（$v_b \geq v_s$），贝叶斯纳什均衡的交易并不必然产生，有效率只是贝叶斯均衡交易的必要条件。

无效率定理显然修正了微观经济学中有关一般均衡的命题，在微观经济学中，只要交易是有效率的，我们就能得到一条契约线，从而保证一般均衡出现。但根据无效率定理，如果经济中存在着私人信息（信息不完全），那么交易有效率并不能保证交易一定发生。为什么会产生有效率而无交易，根本原因就在于效率和激励兼容在绝大多数情况下就像熊掌与鱼肉不可兼得。

第四节　显示原理

在前面讲述拍卖理论的时候,我们已经看到:(1)拍卖机制有多种。(2)明确使用某种拍卖机制后,拍卖中的投标者的类型服从什么样的概率分布,策略空间是不是非常庞大。这些问题十分关键。(3)即使以上问题都明确,也还存在其他的一些可能性,例如,投标者可能还需要支付一定的进入费用,标的物可能有底价,标的物可能卖不出去,卖出去也不总是卖给出价最高的人,等等。这意味着,要找到最优的均衡解,我们必须:首先,依次考虑所有可能的机制,确定每一个机制可能出现的均衡解;其次,比较所有不同机制下的均衡解,找出全局性的最优均衡;最后,要考虑可能出现的一些情况是否会对均衡解造成扭曲。以上这些因素表明,要找到贝氏博弈中的最优均衡解与大海捞针无异。那么,有没有什么办法来最大程度地简化我们所面临的贝氏博弈呢?

一个天真的想法是,如果人人都是讲真话的君子该多好!只要人们把他们的真实类型报告出来,所有的问题都迎刃而解。但在理性人的假设下,一个自然的诘问是,如果报告真实偏好不能带来好处,为什么要真实报告自己的偏好?如果竞标者真实报告自己的偏好那不就意味着要付更多的钱?学生真实报告自己的不努力那不意味着要求老师网开一面没了理由?工人真实报告自己工作偷懒那不意味着加薪泡汤?企业主真实报告收入那不意味着税收翻番的发财梦成为空想?这就给我们一个提示,人们不会主动报告自己的真实偏好,而必须设计一个机制(或制度),在这个机制中参与者不得不报告他的真实偏好。那么能不能设计这样的机制呢?迈尔森(Myerson)1979年提出的显示原理证明这样的机制是存在的。[①]显示原理不仅可运用于拍卖机制(拍卖机制只是

① Myerson, R. 1979:"Incentive Compatibility and the Bargaining Problem", *Econometrica* 47, pp. 61～73.

一个例子),而且可运用在所有的私人信息博弈中,如公共产品配置、外在性解决、垄断监管、道德风险、逆向选择等。这表明显示原理具有非凡的理论意义。这里我们仍然以拍卖机制为例来阐述显示原理。

考虑拍卖者希望设计一个拍卖以使他的期望收益最大化,显然逐一列出卖方可能考虑的不同的拍卖机制,并求最优贝氏均衡是不可能的(例如类型服从不同的概率分布就可能使策略空间无穷大)。这时拍卖者就可以借助显示原理,从两个方面来简化他的问题。

第一,拍卖方可以将其分析集中于下面的博弈:

1. 投标者同时报告(也许是不诚实的)他们自己的类型。把投标者 i 的报告函数定义为 $r_i: T_i \to T_i$,他是从类型空间到类型空间的函数,记作 $r_i(t_i)$,其中 t_i 为参与者 i 的真实类型。如果参与者 i 讲真话,那么 $r_i(t_i)=t_i$。

2. 给定投标者的报告组合 $(r_1, r_2, \cdots, r_n)$,投标者 i 报价 $x_i(r_1, r_2, \cdots, r_n)$ 时,将以 $q_i(r_1, r_2, \cdots, r_n)$ 的概率获得标的物。对于所有可能的报告组合 $(r_1, r_2, \cdots, r_n)$,概率 $\sum_{i=1}^{n} q_i(r_1, \cdots, r_n) \leqslant 1$ 。

我们把参与者报告自己偏好的贝氏博弈称为直接机制。

第二,根据显示原理,存在策略组合 $(r_1(t_1)=t_1, r_2(t_2)=t_2, \cdots, r_n(t_n)=t_n)$ 是直接机制中的一个贝叶斯纳什均衡。这表明,通过设计适当的支付函数和概率函数 $(x_1(r_1, \cdots, r_n), \cdots, x_n(r_1, \cdots, r_n); q_1(r_1, \cdots, r_n), \cdots, q_n(r_1, \cdots, r_n))$,可以使每一个参与者的最优策略是实话实说,即对 $\forall\ t_i \in T_i, r_i(t_i)=t_i$。实话实说成为贝叶斯纳什均衡策略表明,实话实说是参与者的最优策略,每一个参与者都不会偏离它(说假话)。

上述内容的关键是,实话实说要成为直接机制中的贝叶斯纳什均衡策略必须满足激励兼容条件。激励兼容条件通俗地说,就是委托人(拍卖者)希望代理人(竞标者)采取的策略(行动)使得委托人的利益最大,同时也使得代理人本身的利益最大。例如,一位画家(委托人)拍卖他的作品,他当然希望拍卖的价格越高越好,如果他委托一位拍卖师(代理人)帮他拍卖,并规定拍卖师的佣金按拍卖价格的百分比支付,那么拍卖师的策略就满足激励兼容条件,因为拍卖师拍卖的价格越高他

得到的佣金越多，画家得到的收益也越多。一个企业主和工人同样形成了一种委托代理关系，如果企业主付给工人的工资是一个固定数，企业主很难监督工人劳动（存在私人信息），那么工人就可能偷懒，使得企业主利益受损，工人获益（偷懒使效用更大），这时工人的策略（行为）就不满足激励兼容条件。

定理 3.4 显示原理 任何一个贝叶斯博弈都可重新表述为一个满足激励兼容的直接机制，它的任何贝叶斯纳什均衡等价于该直接机制的贝叶斯纳什均衡。

由于任何贝叶斯博弈都对应着一个直接机制，因而我们可以用显示原理来简化对贝叶斯博弈的分析，我们只要分析直接机制就可以了。对于直接机制不存在的特性，不能实现的均衡，相应的贝叶斯博弈也必然不存在这种特性，不能实现这种均衡。例如，根据显示原理我们知道要实现一个对拍卖者最理想的拍卖机制（等同于一个直接机制），它必须满足激励兼容的条件，只有效率是远远不够的，这就（不太严格）证明了我们前面提到的无效率定理——有效率不一定有交易，有效率只是贝氏均衡交易的必要条件。

为了给出显示原理的正式表述及其证明，我们假设存在一个贝叶斯静态博弈 $G_B=\{I,A,T,p,u\}$，其中 $s^*=(s_1^*,\cdots,s_n^*)$ 为它的一个贝叶斯纳什均衡。我们将构建一个由实话实说均衡的直接机制来重新表示 s^*。符合条件的直接机制是一个贝叶斯静态博弈，和 G_B 有相同的类型空间和推断，但其行动空间和收益函数却不一样。由于直接机制中参与者 i 的行动就是（可能是不诚实地）报告他的所属类型，因而直接机制中参与者 i 的行动空间等于类型空间 T_i。新的收益函数则要复杂一些，它们不仅决定于原博弈 G_B，还决定于原博弈的均衡 s^*。问题的关键在于运用 G_B 的均衡解 s^* 来保证实话实说是直接机制的均衡解，具体证明步骤如下：

s^* 是 G_B 的一个贝叶斯纳什均衡，这意味着对每一参与者 i 来说，s_i^* 为 i 对其他参与者策略 $(s_1^*,\cdots,s_{i-1}^*,s_{i+1}^*,\cdots,s_n^*)$ 的最优反应。更为具体地说，对 i 的每一种类型 $t_i\in T_i$，$s_i^*(t_i)$ 都是 i 对给定其他参与者的策略为 $(s_1^*,\cdots,s_{i-1}^*,s_{i+1}^*,\cdots,s_n^*)$ 时，从 A_i 中选择的最优行动。那

么，如果 i 的类型为 t_i，并且我们允许 i 从 A_i 中包含 $s_i^*(t_i)$的一个子集中选择行动（记住是子集），仍然假定其他参与者的策略为$(s_1^*,\cdots,s_{i-1}^*,s_{i-1}^*,\cdots,s_n^*)$，则 i 的最优选择仍为 $s_i^*(t_i)$。选定直接机制中的收益函数时，就是设法使每一参与者面临和上面相同的选择。

确定直接机制中收益函数的方法如下：首先，将新博弈中参与者关于类型的报告 $r=(r_1,\cdots,r_n)$代入原博弈中的均衡策略 s^*，然后将求得的原博弈的均衡行动 $s^*(r)=(s_1^*(r_1),\cdots,s_n^*(r_n))$代入原博弈的收益函数，从而得到直接机制的收益函数。用公式表示，i 的收益函数为：

$$v_i^*(r,t)=u_i[s^*(r),t]$$

其中，$t=(t_1,\cdots,t_n)$。对于这样的收益函数，我们可以理解为存在一个计算机系统，这个计算机系统是按直接机制设计的，它作如下的运算：

(1)所有参与者同时报告自己的类型，并且将它输入计算机中。

(2)计算机根据每一个参与者的报告 r_i，代入原博弈参与者的均衡策略 s_i^*，计算出如果参与者的类型为 r_i 时，他在均衡 s^* 中将会如何行动，也就是 $s_i^*(r_i)$。

(3)计算机根据博弈规则计算出所有参与者在行动 s^* 下各自的收益，并将收益自动记入每一个参与者的账户中。

显然在这个由计算机系统组成的直接机制中，所有参与者惟一的行动就是报告自己的类型（将自己的类型输入计算机），然后参与者得到各自的收益，博弈随之结束。毫无疑问，这个系统是一个典型的贝叶斯静态博弈。它可以表示为 $G_D=\{I,R(T),p,v\}$，其中 $R_i(T)$为报告空间，它实际上等于类型空间 T_i，v_i 为直接机制中的收益函数，它依赖于参与者的行动（报告）和真实类型。

(4)我们证明 $r(t)=t$ 是上述直接机制中的贝叶斯纳什均衡。通过向计算机报告自己的类型 $r_i(t_i)\in T_i$，参与者 i 事实上等于从 A_i 中选择了行动 $s_i^*(r_i)$。如果所有其他参与者都真实报告自己的类型 $r_{-i}(t_{-i})=t_{-i}$，则等于选择了策略 $s_{-i}^*=(s_1^*,\cdots,s_{i-1}^*,s_{i+1}^*,\cdots,s_n^*)$。而前面我们已经知道，$s^*$ 是原博弈的贝叶斯纳什均衡，如果其他参与者选择了这样的策略，当 i 的类型为 t_i 时，参与者 i 的最优行动只能是

$s_i^*(t_i)$。这意味着，如果其他参与者真实报告自己的类型，则参与者 i 的最优策略也只能是真实报告自己的类型。现在设 i 为参与者集合 I 中的任意一个值，即$\forall i \in$ I，上述结论同样成立，即实话实说是直接机制中的一个贝叶斯纳什均衡。更为正式地说，在静态贝叶斯博弈 $G_D=\{I, R(T), p, v\}$ 中，对$\forall t_i \in T_i$，每一参与者 i 选择实话实说策略 $r_i(t_i)=t_i$ 构成一个贝叶斯纳什均衡。

显示原理最重要的理论意义就在于它表明，在广泛的社会选择中存在着不可能性定理，著名的阿罗不可能性定理只不过是不可能性定理在社会公共选择下的特例。所谓不可能性定理是指，在每一个人都平等自由的条件下，各自根据自身的偏好理性地作出选择，然后通过我们所能想到的任何民主制度来合成个人选择以形成统一的社会选择，那么这个社会选择必然是无效率的，也就是效率和民主制度是冲突的。正像显示原理所揭示的任何存在私人信息的制度（直接机制和间接机制）都可表示成说真话做实事的直接机制，然而说真话做实事并不能给经济人带来利益的最大化，要让人们讲真话做实事就必须给予激励，而要实现这种激励就必然以牺牲效率为代价。

在什么情况下人类能摆脱不可能性定理的羁绊？只有在类似于完全竞争这样的制度结构中不可能性定理才会失效。完全竞争的一个重要条件是市场参与者的人数趋于无穷大，从这个意义上来讲，普选＋直选是民主制度的最高形式。美国 2000 年总统大选出现的由法官来决定总统的争端实际上只不过是不可能性定理在社会政治选择中时不时体现自己存在的一种表现而已。为什么没有一个国家的民主制度能够达到普选＋直选？其根本原因在于，普选＋直选存在着巨大的交易费用，而交易费用的产生与信息不完全存在着极大的联系，而在完全竞争条件下信息是完全的，交易费用为零，而这在现实中不过是幻想而已。这使我们不得不想起在公共地悲剧中哲学家怀海特所说的话：问题的解决不在于技术而在于精神。人类物质技术的不断进步并不是人类文明进步的充分条件，人类各种难解之症的解决与其说依赖于技术，不如说依赖于人类崇高的精神和思想。

复习思考题

1. 什么是贝叶斯纳什均衡？为什么贝叶斯法则等同于参与者理性？豪尔绍尼转换的意义是什么？

2. 请你简要证明，为什么说完全信息静态博弈中的混合策略均衡是贝叶斯纳什均衡的特例？

3. 考虑一个密封式一级价格拍卖，其中竞标者对标的物的估价相互独立，并服从区间[0,1]上的均匀分布。证明如果有 n 个竞标者，则以各自估价的 $(n-1)/n$ 倍作为投标价格是这一拍卖的一个对称贝叶斯纳什均衡。

第二部分　动态博弈

动态博弈与静态博弈最大的不同，就在于静态博弈只有一局，即所有参与者同时出招之后博弈结束，而动态博弈则不止一局，且参与者的策略有先后之分。在现实生活中，大多数的情况都是动态的。例如下棋、打牌等游戏，足球、篮球等体育比赛，政治谈判、军事对抗等冲突，都属于动态博弈。动态博弈比静态博弈复杂得多，一个简单的理由是静态博弈是动态博弈的特例，动态博弈具有一些静态博弈所不具有的特性。严格来讲，在静态博弈中，不存在过去也不存在将来，而在动态博弈中，参与者的将来决策，必须取决于现在所做的决策，而现在的决策又必须考虑将来可能出现的情况，即人们常说的"人无远虑，必有近忧"。在这种情况下，策略空间会变得非常巨大，给求解带来极大的困难。例如，人们至今也很难编出较"聪明"的围棋程序，原因就在于当考虑十步、二十步之后的情况时，策略空间大得没有哪台计算机能够算出来。另一个困难就是所谓的时间一致性(Time consistency)问题。所谓时间一致性是指如果策略在博弈的局部是最优的，它也应当在整个博弈中是最优的。在现实生活中，绝大多数情况下，人们以为现在的决策是最优的，但可能在未来发现这只不过是错误的决策。而时间一致性要求我们在计算动态博弈时，绝对不能犯这种错误。

第四章讨论的动态博弈是相对简单的完全信息动态博弈，其核心概念是子博弈完美纳什均衡。第五章则讨论较复杂的非完全信息动态博弈，非完全信息动态博弈中的均衡概念是博弈论中的均衡概念的集大成者。

第四章

完全信息动态博弈

对于完全信息动态博弈，一个自然的想法是，我们应该能把纳什均衡的概念运用到完全信息动态博弈中，因为是完全信息，所以我们能够计算出博弈的每一个阶段的纳什均衡，那么每一个阶段的纳什均衡组成的策略组合有可能就是整个动态博弈中的均衡。这种想法有其合理性，例如一个人坐火车从北京到上海，其中必须经过济南，那么从北京到济南最短的路线与从济南到上海最短的路线之和一定是北京到上海的最短路线。但动态博弈远比我们上面的例子复杂，简单地套用纳什均衡会产生许多不论从直观上还是从逻辑上都很不合理的结果，其中最重要的问题就是所谓的不可置信。

在这一章中，我们首先对博弈的扩展式表达给出一个完整的定义，为动态博弈的分

析打下基础;其次,我们会从最简单的动态博弈入手,逐步深入;最后分析具有无穷次的重复博弈,并推导出无名氏定理,从而为人类合作逃脱囚徒困境提供有力的说明。在本章的最后,给出子博弈完美纳什均衡的完整定义。

第一节　博弈表达的扩展式

博弈扩展式是针对博弈树(Game tree)这个工具来定义的。在前面讲述静态博弈时,我们已经反复用到了这个工具,因而对博弈树并不陌生,但我们并没有给出严格的定义,原因之一是要给出扩展式的完整定义只有在动态非完全博弈的情况下才有可能。我们还是从例子开始。

例 4.1　囚徒困境第三版。

囚徒困境第三版与原始版本惟一不同的地方是,囚徒 1 首先行动(招供或沉默),然后被囚徒 2 所观察到,囚徒 2 再相应地采取行动。显然这是一个最简单的完全信息动态博弈,用博弈树可表示成图 4.1。

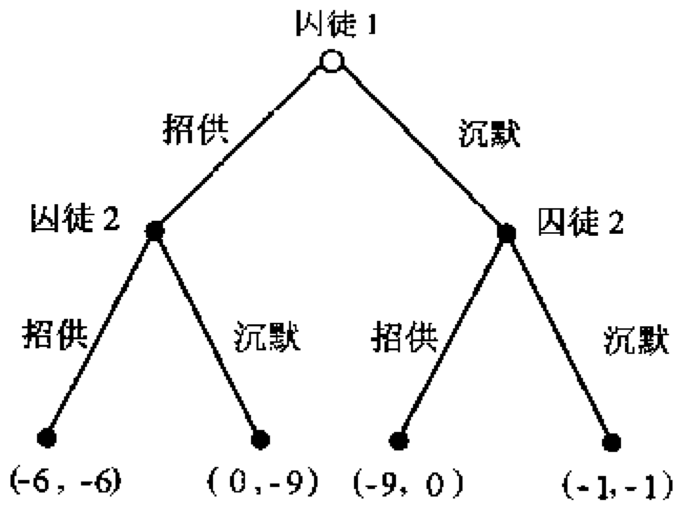


图 4.1　囚徒困境第三版

这个博弈开始于一个起始决策点,通常用空心圆表示,在这一点上囚徒 1 作出行动,要么招供要么沉默,它有两个分枝分别表示招供和沉

默，在枝的终点是囚徒 2 的决策点，通常用实心圆表示，在这一个点上囚徒 2 同样有两个行动可以采取。由于两个点是分开的，这表明囚徒 2 知道自己在做决策时到底位于哪一个点，即囚徒 2 能够观察到囚徒 1 所做的选择。囚徒 2 行动之后，博弈就进行到结束点。在每一个结束点，对应着囚徒 1 和囚徒 2 一系列行动后导致的收益。这表明博弈树是全排序的，即从博弈树的起点到博弈树中的任何一点只有惟一一条路径。只要读者细心观察真实的树，就会发现真实的树也是全排序的，由于它与真实的树一样，因而被称为博弈树。

博弈的扩展式实际上就是对博弈树的节点（Node）和枝（Branch）作出严格的规范和定义，因而博弈的扩展式与博弈树是同义语。与博弈的基本式相比，博弈的扩展式特别善于描述复杂的博弈，博弈越复杂，博弈扩展式的优点就越突出。比如，下象棋这个博弈用基本式来表述就极其繁琐，而用扩展式来表述却十分简单。下象棋通常需要两个参与者，我们定义为红方和黑方，分别用 R 和 B 来表示，红方先走，黑方后走，这是一个典型的完全信息动态博弈。象棋这个游戏可表示成图4.2。其他的内容我们省略掉了。

在囚徒困境第三版和象棋游戏这两个博弈中，当每一个参与者行动时，他都能观察到他的对手先前所有的行动。这种博弈我们定义为完美信息（Perfect information）博弈。所谓完美信息是指博弈中的每一个信息集（Information set）只包含惟一一个节点。信息集实际上是参与者决策节点集合的一个子集。信息集的定义是当博弈进行到某个阶段，轮到某个参与者行动时，他不知道自己位于哪一个节点，不能加以区分的节点的集合就称为信息集。参与者之所以不能对信息集中的节点进行区分，原因在于他对前面发生的其他参与者的（有些或全部）行动未能观察到。现在我们看图 4.3。它是原版的囚徒困境。

在图 4.3 中，我们看到当博弈进行到囚徒 2 行动时，由于囚徒 2 未能观察到此前囚徒 1 的行动，所以囚徒 2 不能对节点 x_2 和 x_3 作出区分，也就是他不知道囚徒 1 到底选择的是什么策略。节点 x_2、x_3 的集合 $\{x_2, x_3\}$ 即是囚徒 2 的一个信息集，集合 $\{x_1\}$ 即是囚徒 1 的信息集。习惯上，我们通常把位于同一个信息集的节点用虚线连接起来或者用圆

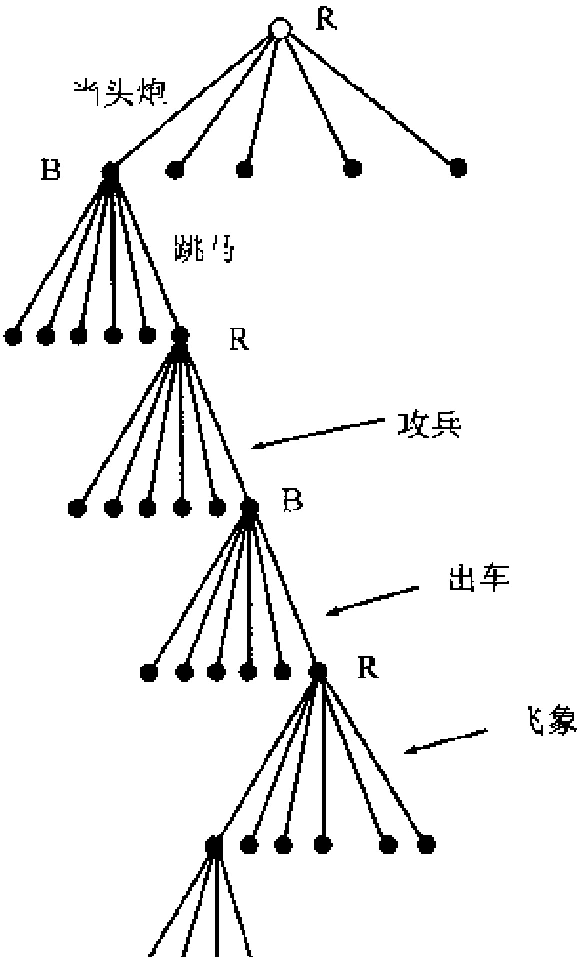


图 4.2　象棋的扩展式

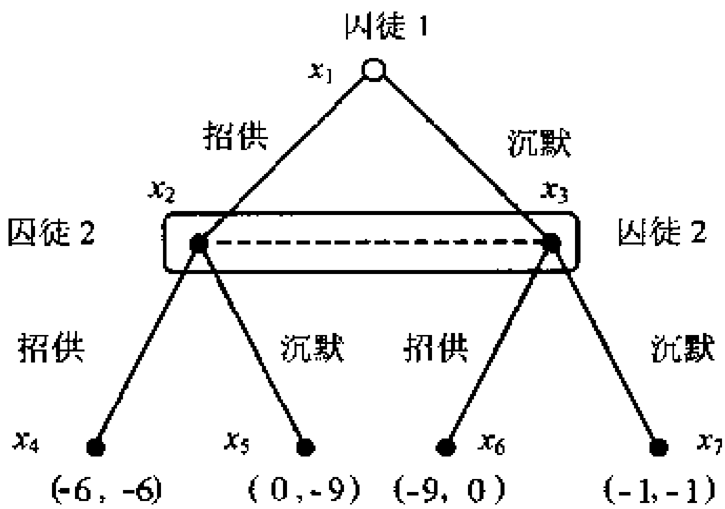


图 4.3　囚徒困境的扩展式

圈圈起来，只包含一个节点的信息集就称为单点信息集(Singleton information set)。

首先，使用信息集可以清楚地描绘出当参与者行动时，他所处的“环境”和所有可能掌握的信息，因而能揭示出博弈的信息特征。其次，信息集的使用还可以使静态博弈转化为动态博弈，在博弈的基本式和扩展式之间架设了一道桥梁，博弈的扩展式成为能够描述所有博弈类型的有力工具。

对于信息集有一些自然而严格的规定：(1)在任意一个信息集中的每一个节点上，参与者一定有相同的行动集合。原因非常简单，如果一个信息集中的两个节点包含的行动不一样，由于参与者在不同的节点具有不同的行动，因而他一定能对这两个节点作出区分。这就与信息集的定义相矛盾。(2)信息集不得违背完美回忆(Perfect recall)的假定。完美回忆假定当博弈中的参与者行动时，他对过去发生的一切情况都记得清清楚楚。显然完美回忆在现实生活中可能并不满足，因而我们把它看作一种假定，但完美回忆在现实生活中也普遍存在，例如，打过桥牌的人就知道桥牌游戏就是一个典型的完美回忆。违背完美回忆的情况如图 4.4 所示。

在图 4.4 中，参与者 2 忘记了他曾经知道的参与者 1 如何行动的信息。如果一个博弈不满足完美回忆的要求，实际上我们就不可能对该博弈进行分析，因为忘记，所以一切后果都可能出现。在博弈论中，完美回忆和博弈的结构是所有参与者的公共信息，是每一个博弈都必须满足的条件。

定义 4.1　一个博弈是完美信息博弈，它的每一个信息集都是单点信息集，否则就是非完美信息。

博弈的扩展式除可以用博弈树表示之外，还可以通过数学来严格定义。

定义 4.2　一个博弈的扩展式记为 $\Gamma=\{I, X, A, p(\cdot), k(\cdot), a(\cdot), H, \rho, u\}$，其中：

(1) X 为该博弈所有节点的集合，A 为博弈所有行动的集合，$I=\{1,\cdots,n\}$ 为参与者集合。

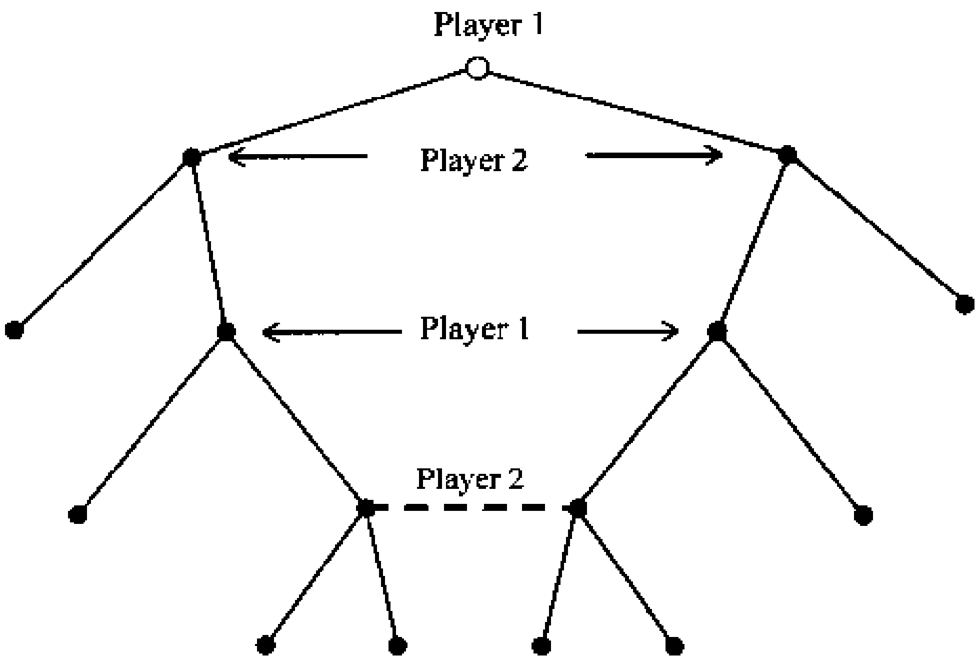


图 4.4　完美回忆的违背

(2)函数 $p: X \to \{X \cup \varnothing\}$ 表示每一个节点 x 的最近前列点(Immediate predecessor)。$s(x)=p^{-1}(x)$就定义为节点 x 的最近后列点(Immediate successor)。我们用 $x>y$ 来表示 x 是 y 的前列点。当 x 为起始点(Initial node),$p(x)$就为空集,通常用 x_0 表示。除起始点外,任何节点的最近前列点必然是惟一的。这意味着博弈中的节点是全排序的,通俗地说就是从起始点到博弈中的任何一节点的路径是惟一的;反过来,从任何一节点到博弈的起始点的路径也是惟一的。这说明任意两个节点 x_i 和 x_j(要么无关)要么 $x_i>x_j$,要么 $x_j>x_i$,不能两者都成立。通过不断重复 p 和 s,我们就能找出博弈所有的路径。$T=\{x\in X; s(x)=\varnothing\}$ 为所有结束点(Terminal node)的集合,$X|T$ 为所有决策点(Decision node)的集合。例如,在图 4.3 中,x_1 为 x_2 的最近前列点,即 $p(x_2)=x_1$,x_1 和 x_2 不能互为前列点,x_4 是 x_2 的最近后列点,即 $s(x_2)=x_4$,但 x_1 不是 x_4 的最近前列点,而是 x_4 的前列点,即 $x_1>x_4$。

(3)函数 $k: X|T\to I$。$k(x)$表示在 x 决策点作决策的是哪一个参与者。例如在图 4.3 中,$k(x_1)=$囚徒 1。

(4)函数 $a: X|T\to A$。$a(x)$代表在决策点 x 上,所有可供参与者选择的行动集合。例如图 4.3 中,$a(x_1)=\{$招供,沉默$\}$。$A_i=\{a(x)|k(x)$

$=i,x\in X|T\}$为参与者 i 的所有可用行动的集合。

(5)函数 $h^i(x)$代表参与者 i 在决策点 x 上的信息集。H^i 为参与者 i 的所有信息集的集合，即 $H^i=\{h^i(x)|k(x)=i,x\in X|T\}$。所有信息集的集合为 $H=\{H^1,H^2,\cdots,H^n\}$。H 实际上是对 $X|T$ 的一个分割。[1]由于信息集是参与者不能分辨的决策点集合，所以可用信息集来代替决策点定义函数 a 和 k，即 $a(x)=a(h),k(x)=k(h)$。

(6)函数 $\rho:H_0\times A\to[0,1]$表示自然行动的概率分布。H_0 为起始点的信息集。例如在图 3.4 中 ρ(讲道义)$=\theta$，ρ(不讲道义)$=1-\theta$。

(7)收益函数集合 $u=\{u_1,\cdots,u_n\}$表示当到达某个结束点时，所有参与者获得的收益，其中 $u_i:T\to R$。如果博弈的结果是随机实现的，那么收益函数就是 v－N－M 收益函数。

我们曾定义策略是参与者的完全行动计划。在静态博弈(以后我们称为基本式博弈)下策略等于参与者的行动或是行动的一个概率分布，它实际上并没有真正体现“完全”这个重要特征，只有在动态博弈(扩展式博弈)中策略的这一定义才能得到真正的体现。

一个参与者在扩展式博弈 Γ 中的策略，就是参与者每次轮到他行动时，他要给出如何行动的一个组合。严格的定义如下。

定义 4.3 在扩展式博弈 $\Gamma=\{I,X,A,p(\cdot),k(\cdot),a(\cdot),H,\rho,u\}$中，对于参与者 $i\in I$ 而言，一个信息集 $h^i(x)\in H^i$ 和行动 $a_j\in a(h^i)$决定的一个概率 $\lambda_i(a_j,h)\geqslant 0$ 就称为参与者 i 的一个行为策略(Behavior strategy)，其中 $\sum_{a_j\in a(h^i)}\lambda_i(a_j,h)=1$。

显然行为策略实际上就是基本式博弈中的混合策略当中的一项。在扩展式博弈中，参与者的策略实际上就是参与者在博弈的每一阶段行为策略的集合。

定义 4.4 在扩展式博弈$\Gamma=\{I,X,A,p(\cdot),k(\cdot),a(\cdot),H,\rho,u\}$

[1] 所谓分割就是设 Y 为一个集合，$Y_1,Y_2,\cdots,Y_n$ 为集合 Y 的子集，如果 $Y_1\cap Y_2$，$Y_2\cap Y_3$ 等任意两两相交为空集，$Y_1\cup Y_2\cup\cdots\cup Y_n=Y$，那么就称 $Y_1,Y_2,\cdots,Y_n$ 为集合 Y 的一个分割。

中，参与者 i 的一个阶段策略为一个函数 $\sigma_i: \{h\in H^i | k(h)=i\}\to\{\Delta a(h)|k(h)=i\}$，其中 $\Delta a(h)$ 为信息集 h 下所有行动的概率分布。$\sigma_i=\{\sigma_i(h)|all\ h\in H^i\}$ 我们称为参与者的一个完整策略，简称策略。

显然对于任一信息集 $h\in H^i$，$\sigma_i(h)$ 是 $a(h)$ 上的一个概率分布，它等于 $(\lambda_i(a_1,h),\cdots,\lambda_i(a_k,h))$，其中信息集 h 包含的行动为 $\{a_1,\cdots,a_k\}$。实际上根据定义，一个完整的 $\sigma_i(h)$ 才能被看作是一个策略，但经济学家约定俗成，把 $\sigma_i(h)$ 和 σ_i 都称做策略。

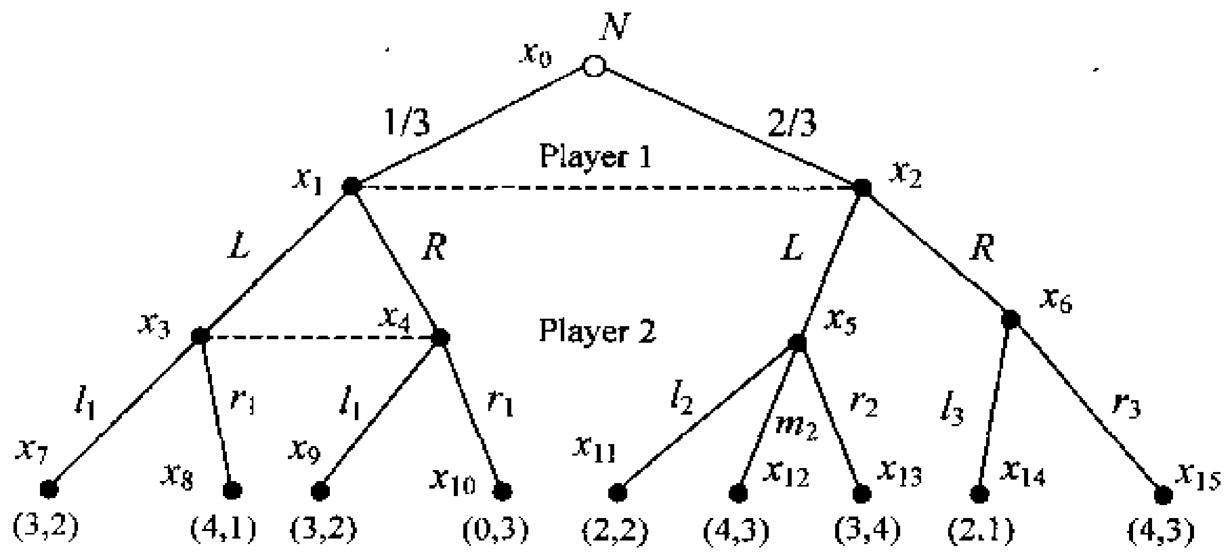


图 4.5 非完全信息博弈

现在我们举例说明，如图 4.5 所示。$\sigma_1(\{x_1,x_2\})=(\frac{1}{2}L,\frac{1}{2}R)$代表参与者 1 的一个策略，它的意思是参与者 1 在信息集$\{x_1,x_2\}$中用 1/2 的概率出 L，1/2 的概率出 R，其中 1/2 概率下 L 即为一个行为策略。又如，$\sigma_2(\{x_6\})=(\frac{3}{5}l_3,\frac{2}{5}r_3)$；$\sigma_7(\{x_5\})=(\frac{1}{3}l_2,\frac{1}{3}m_2,\frac{1}{3}r_2)$；$\sigma_2(\{x_3,x_4\})=(\frac{1}{3}l_1,\frac{2}{3}r_1)$则为参与者 2 的一个策略。它的意思是表示参与者 2 走到 x_6 时，他的策略为 3/5 的概率出 l_3，2/5 的概率出 r_3；当走到 x_5 时，他的策略为 1/3 的概率出 l_2，1/3的概率出 m_2，1/3 的概率出 r_2；当走到信息集$\{x_3,x_4\}$时，他的策略为 1/3 的概率出 l_1，2/3 的概率出 r_1。定义 $\sigma=(\sigma_1,\cdots,\sigma_n)$为一个策略组合，$\sum_i$ 为参与者 i 所有可能策略的

集合，令 $\sum = (\sum_1, \cdots, \sum_n)$ 为所有参与者可能策略组合的集合。对任意一个 $\sigma \in \sum$，第 i 个参与者的收益为：

$$u_i(\sigma) = \sum_{z \in T} \pi^{\sigma,\rho}(z) u_i(z)$$

其中，$\pi^{\sigma,\rho}(z)$ 代表在某个策略组合 σ 下，到达某个终点 $z \in T$ 的概率。

例如，利用上例的 $\sigma = (\sigma_1, \sigma_2)$，可算出 $\pi^{\sigma,\rho}(x_7) = \pi^{\sigma,\rho}(x_9) = \frac{1}{3} \times \frac{1}{2} \times \frac{1}{3} = \frac{1}{18}$，$\pi^{\sigma,\rho}(x_{11}) = \pi^{\sigma,\rho}(x_{12}) = \pi^{\sigma,\rho}(x_{13}) = \frac{2}{18}$，$\pi^{\sigma,\rho}(x_8) = \pi^{\sigma,\rho}(x_{10}) = \frac{2}{18}$，$\pi^{\sigma,\rho}(x_{14}) = \frac{1}{5}$，$\pi^{\sigma,\rho}(x_{15}) = \frac{2}{15}$，因此 $u_1(\sigma) = \frac{122}{45}$，$u_2(\sigma) = \frac{102}{45}$。

第二节　子博弈完美纳什均衡

由于动态博弈可以看作是由一串静态博弈组成的，一个自然的想法是纳什均衡的概念能否在扩展式博弈中加以运用呢？答案是肯定的，但也带了一些不理想的结果。由于纳什均衡的概念在动态博弈下要求过于宽松，因而一些不合理的均衡在动态博弈中出现了，这些不合理的均衡大部分都是不可置信的(Incredible)。下面我们将从纳什均衡入手引出问题，最后给出子博弈完美纳什均衡(Subgame Perfect Nash Equilibrium)的概念。为了方便，以后简称完美均衡，简记SPNE。

一、纳什均衡

一个自然的想法是，将基本式博弈中的纳什均衡概念原封不动地套用到扩展式博弈中，并用它来求解动态博弈中的均衡。这样做可能存在问题，但也为我们找到动态博弈的均衡开了一个好头。由于动态博弈的复杂性随着博弈结构的增加呈几何倍数增加，因而开个好头非常重要，

其意义可以用诗句“山重水复疑无路，柳暗花明又一村”来加以形容。

根据纳什均衡的定义，所有参与者的均衡策略都应是其他对手的最优反应，也即在扩展式博弈 Γ 中，纳什均衡是一个策略组合 $\sigma^* \in \sum$，使得 $u_i(\sigma_i^*, \sigma_{-i}^*) \geqslant u_i(\sigma_i, \sigma_{-i}^*)$，对所有的 $\sigma_i \in \sum_i, i \in I$ 成立。

在图 4.6 所表示的这个动态博弈中，哪一个策略组合会是博弈的均衡解呢？

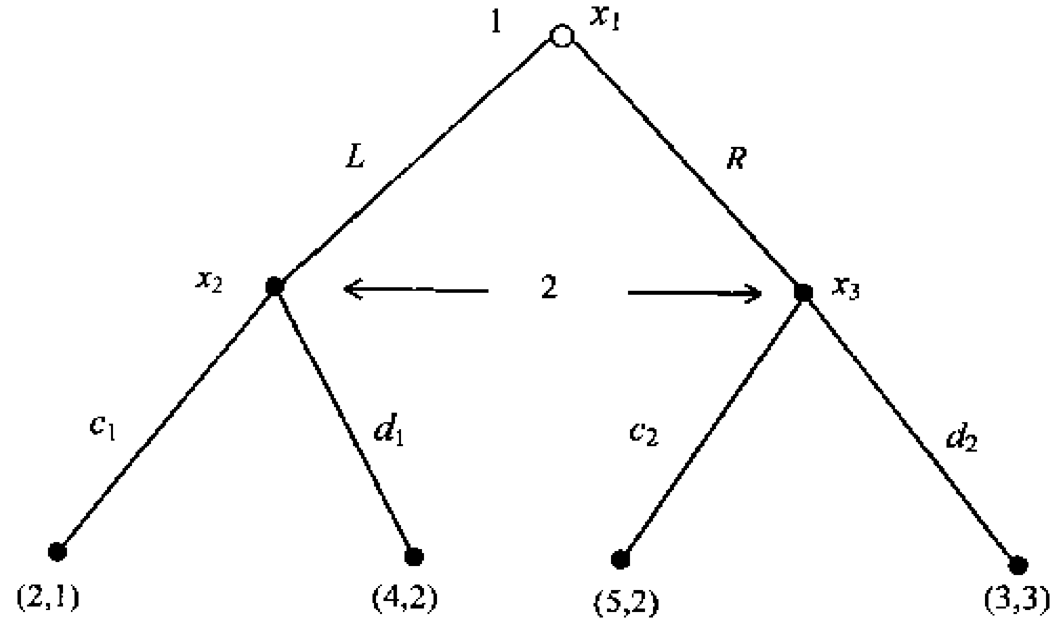


图 4.6　完美信息博弈

根据纳什均衡的定义，当参与者 1 的策略为 $\sigma_1^*(x_1)=R$，参与者 2 的策略为 $\{\sigma_2^*(x_2)=c_1, \sigma_2^*(x_3)=d_2\}$ 时，策略组合 $\sigma^*=(\sigma_1^*, \sigma_2^*)$ 为一个纳什均衡。

证明：首先给定参与者 2 的策略为 σ_2^*，参与者 1 选择一个策略 $\sigma_1=(tL,(1-t)R)$ 使得他的期望收益最大，即：

$$\begin{aligned} Eu_1(\sigma_1, \sigma_2^*) &= tu_1(L, \sigma_2^*) + (1-t)u_1(R, \sigma_2^*) \\ &= tu_1(L, c_1) + (1-t)u_1(R, d_2) \\ &= t \cdot 2 + (1-t) \cdot 3 = 3 - t \end{aligned}$$

参与者最优的策略是令 $t=0$，也就是出 R，而不出 L，这实际上就是 σ_1^*。

其次考察参与者 2 的最优策略。给定参与者 1 的策略为 σ_1^*，参与者 2 的一个策略 $\sigma_2=\{\sigma_2(\{x_2\})=(sc_1,(1-s)d_1),\sigma_2(\{x_3\})=(vc_2,(1-v)d_2)\}$ 应使期望收益最大，即：

$$
\begin{aligned}
Eu_2(\sigma_1^*,\sigma_2)&=\pi^{\sigma,\rho}(x_2)u_2[L,(sc_1,(1-s)d_2)]+\pi^{\sigma,\rho}(x_3)u_2[R,\\
&\quad (vc_2,(1-v)d_2)]\\
&=0\times[su_2(L,c_1)+(1-s)u_2(L,d_1)]+1\times[vu_2\\
&\quad (R,c_2)+(1-v)u_2(R,d_2)]\\
&=v\times 2+(1-v)\times 3=3-v
\end{aligned}
$$

参与者 2 的最优策略是让 $v=0$，即在节点 x_3 出 d_2。至于 s 的值，由于参与者 1 选择 L 的可能性为零，即 $\sigma_1(L)=0$，所以参与者 2 在信息集 x_2 下无论如何选择都不会影响其期望收益的值。因此 $\sigma_2^*=(c_1,d_2)$ 是参与者 2 面对 $\sigma_1^*=R$ 的最优反应。综上所述，$(\sigma_1^*=R,\sigma_2=(c_1,d_2))$ 就是扩展式博弈 Γ 的一个纳什均衡解。

但只要我们稍加细想，就会发现这个所谓的纳什均衡包含着不合理的因素，他在现实中根本不会出现。其原因就在于参与者 2 在信息集 x_2“威胁”出 c_1 是一个难以置信的承诺。之所以造成这种情况，是由于在参与者 1 的策略为 R 下，博弈进行到 x_2 的可能性为零，因此参与者 2 在信息集 x_2 下的任何行动都不会对他的最终收益造成影响。这意味着，在纳什均衡中，参与者 2 在 x_2 可以任意选择(像一个非理性的人一样)。这就可能导致这样的结论：参与者 2 的策略是一个前后不一致的(Time inconsistent)的策略。一个前后不一致的策略肯定不会是一个最优的策略。我们可以这样理解参与者 2 的行动：参与者 2 之所以威胁当参与者 1 出 L 时，他要选择 c_1，目的在于通过威胁使参与者 1 选择有利于参与者 2 的 R，因为在参与者 1 选择 R 时，参与者 2 通过选择 d_2，能得到 3 的报酬，明显好于当参与者 1 选 L，参与者 2 选 d_1 时的收益 2。但我们要问的是，如果参与者 1 不顾参与者 2 的威胁而选择了 L，参与者 2 可能会选择 σ_2^* 所指定的 c_1 吗？在参与者 2 为理性是公共信息的条件下，参与者 2 选择 c_1 的报酬为 1，而选择 d_1 的报酬为 2。由于 d_1 要优于 c_1，因而我们没有理由相信参与者 2 会实现他的承诺，也就是

说，参与者 2 的策略 σ_2^* 是一个难以置信的威胁。如果威胁成真，σ_2^* 就是一个前后不一致的策略，因为参与者 2 事前是理性的，但在博弈进行到 x_2 时，他却成了一个非理性的人（选择了 c_1，而不是 d_1）。上述纳什均衡所存在的问题，是把静态博弈中的纳什均衡概念原封不动套用到扩展式博弈时必然会出现的。因为一个纳什均衡只要求在博弈的总体上，参与者的策略须为均衡，而对博弈进行到某个部分时是否为均衡没有要求，这就可能导致总体和局部的冲突，产生一些不合理的纳什均衡。用数学来分析，这种情形之所以出现，是因为有些信息集在原先的策略下走到的概率为 0，因此在这些信息集所做的事情，对每个参与者的收益都没有影响，因而参与者在此可以任意选择行动而丝毫不会影响到原先纳什均衡的收益。例如在图 4.6 中，参与者 2 不论在 x_2 选择什么行动，他的收益都一样（因为在 σ^* 下，走到 x_2 的可能性为 0，而 0 乘以任何数都为 0），所以他可以在 x_2 下任意选择策略，但这明显与参与者的理性相冲突。假如参与者 1 选择了 L 而不是 R，这时参与者 2 就会走到 x_2，那么他还会选择 c_1 以实现他的承诺吗？答案是否定的，因为选择 d_1 的收益为 2，要大于选择 c_1 的收益 1，因而参与者 2 最理性的选择是 d_1 而不是 c_1。一旦如此，参与者 1 发现他选择 L 而不是选择 R，那么他的收益会从原先的 3 增加到 4。参与者 1 改变策略当属明智之举。明确了这一点我们就会发现原来的策略组合（$\sigma_1^*=R$，$\sigma_2=(c_1,d_2)$）不再是纳什均衡，真正的纳什均衡应当是参与者 1 不理会参与者 2 的威胁，选择 L，参与者 2 在 x_2 选择 d_1，在 x_3 选择 d_2，因此真正的纳什均衡为（$\sigma_1^*=L$，$\sigma_2=(d_1,d_2)$）。

综上所述，我们看到在一个存在现在与将来的动态博弈中，需要一个比纳什均衡更强的均衡概念。简言之，在动态博弈中的均衡概念，不但要在整个博弈中是均衡的，而且要在他的任何局部也是均衡的；不但现在是均衡的，过去、将来也仍然是均衡的。只有在这种要求下，博弈的参与者才能实现策略的前后一致性。这就引出了子博弈完美纳什均衡（SPNE）的概念。

二、子博弈完美纳什均衡

子博弈通俗地说就是动态博弈的局部，只要是一个单点信息集，它后面的所有节点（信息集和结束点）连同它本身就构成了一个子博弈。如果一个动态博弈中的纳什均衡在所有该博弈的子博弈中仍然是纳什均衡，那么我们前面遇到的不可置信威胁或是策略的前后不一致性就不会出现，这一思想最早由经济学家塞尔腾（Selten）于 1965 年提出。也主要是由于这个贡献，塞尔腾与纳什、豪尔绍尼共同获得了 1994 年的诺贝尔经济学奖。[1]

定义 4.5 一个扩展式博弈 $\Gamma=\{I,X,A,p(\cdot),k(\cdot),a(\cdot),H,\rho,u\}$ 中的**子博弈**是该博弈的一个子集，其具有下列特征：

（1）它必须从单点信息集（包含惟一一个决策点）开始，包含该决策点的所有后列点，并且只包含以上这些节点。

（2）假如决策点 x 是子博弈的一个决策点，那么每一个 $x'\in h(x)$ 同样是该子博弈的决策点。也就是说，如果子博弈开始于节点 y，那么 $y>x$ 就意味着对所有 $x'\in h(x)$ 而言，$y>x'$，即子博弈不能"分割"任何信息集。

在图 4.7 中，如果我们把整个博弈也看作是一个子博弈（通常我们都这样处理），那么它有五个子博弈。x_1 为第一个子博弈，x_2 为第二个子博弈，x_3 为第三个子博弈，x_4 为第四个子博弈，加上原来的博弈，总共五个子博弈。

根据子博弈的定义，在图 4.5 中，只存在三个子博弈，即该博弈本身和 x_5、x_6 两个子博弈，尽管它看起来节点和分支比图 4.7 要多，原因就在于子博弈不能分割信息集。由于图 4.5 中包含着两个非单点信息集，因而使得图 4.5 中的博弈只包含三个子博弈。图 4.8 则给出了违背子博弈的情况。图 4.8 与图 4.7 实际上拥有一样的节点和分支，但只有

[1] Selten, R. 1965:"Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit", *Zeitschrift fur Gesamte Staatswissenschaft* 121: 301～324. and 1975:"Re examination of the Perfectness Concept for Equilibrium Points in Extensive Games", *International Journal of Game* 4:25～55.

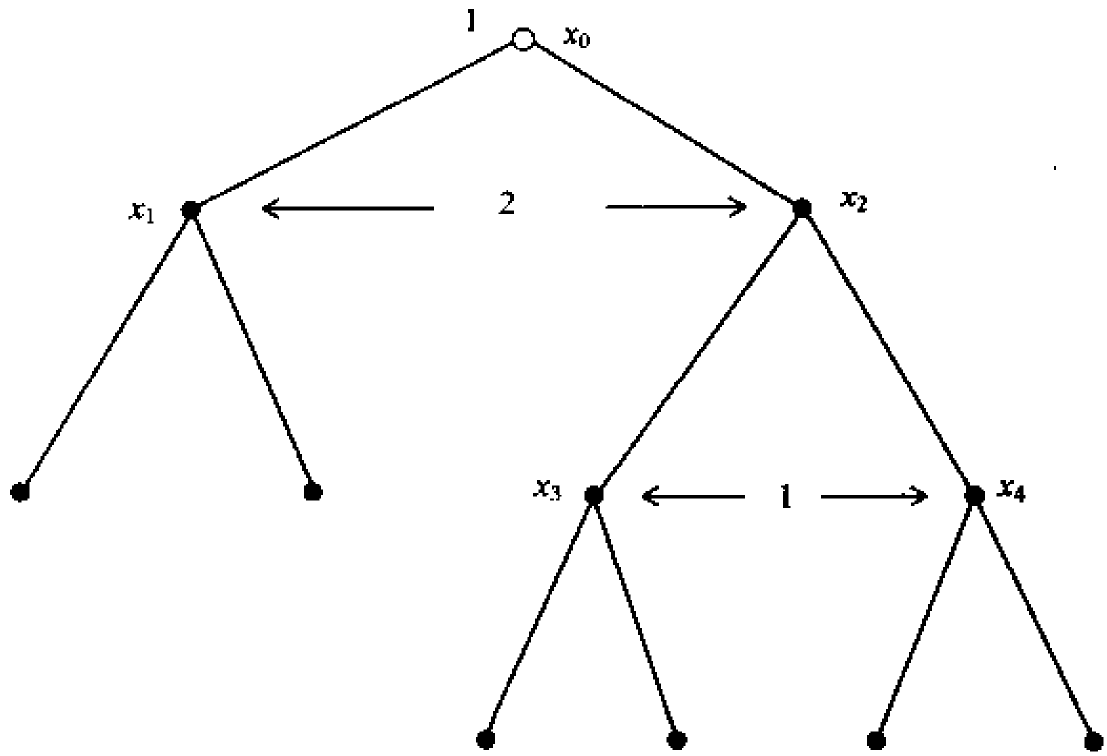


图 4.7　存在着 5 个子博弈

三个子博弈存在。

在图 4.8 中，虚线方框中的部分不是子博弈，因为它的起始点不是惟一的；而虚线椭圆中的部分也不是子博弈，因为它分割了一个信息集 $\{x_3, x_4\}$。

子博弈的核心就在于它本身就是一个博弈，因而纳什均衡的概念就可以运用到子博弈之中，这就使我们在前面建立起来的理论可以运用到动态博弈中，而没有必要另起炉灶。如果一个策略组合在所有的子博弈中都是纳什均衡，显然不可置信威胁和策略上的前后不一致都能加以克服，从而收到了一石二鸟之效。写到这我们不得不叹服理论之美，理论之妙，人类智慧之伟大。

定义 4.6　在扩展式博弈 $\Gamma=\{I, X, A, p(\cdot), k(\cdot), a(\cdot), H, \rho, u\}$ 中，如果一个策略组合 $\sigma^*=(\sigma_1^*, \cdots, \sigma_n^*)$ 在所有该博弈的子博弈中都是纳什均衡，那么我们就称策略组合 $\sigma^*=(\sigma_1^*, \cdots, \sigma_n^*)$ 为子博弈完美纳什均衡，简称完美均衡，简记为 SPNE。

对于一个完全信息动态博弈我们如何去解它的完美均衡呢？一个

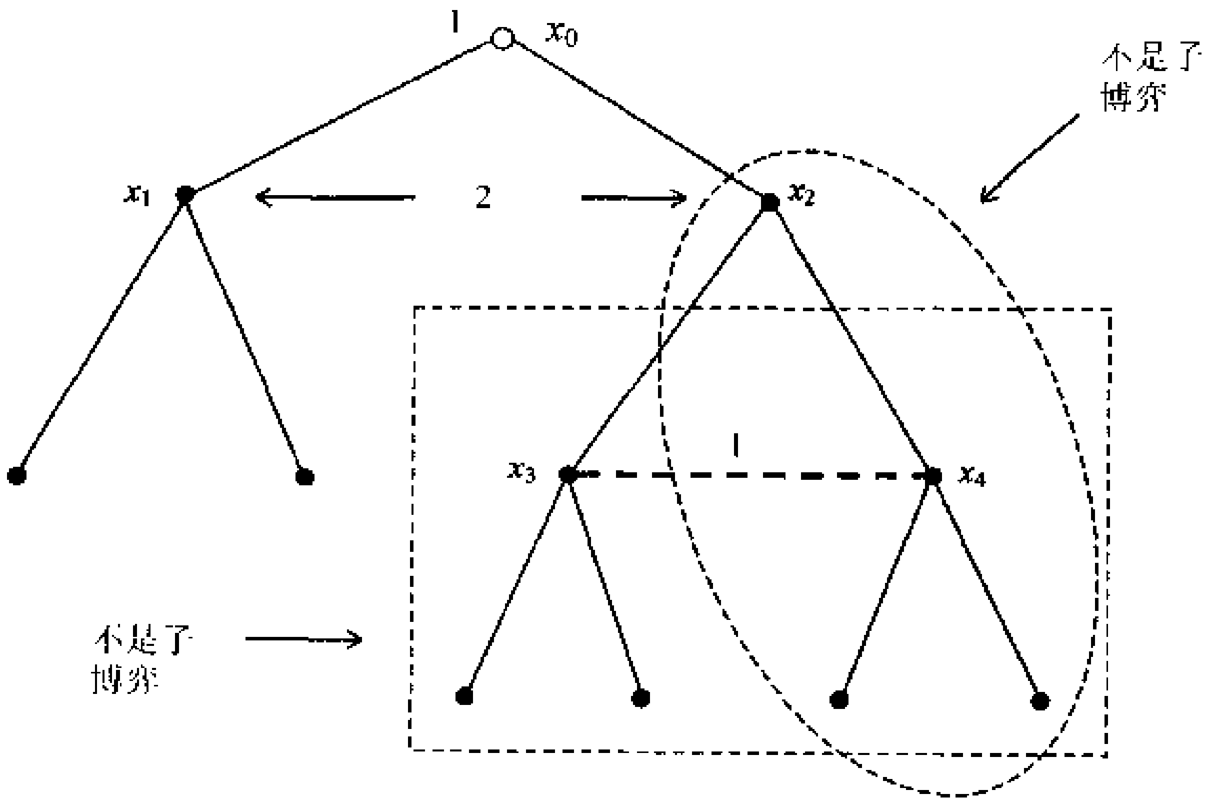


图 4.8　不是子博弈的例子

一般的方法就是所谓的逆推法(Backward induction procedure)。由于博弈的结构是全排序的，因而运用逆推法将能够走遍所有的子博弈，不会有遗漏。

定义 4.7　所谓逆推法是指如下程序：

第一步，从扩展式博弈(博弈树)的终点开始，找到每一个最后子博弈(它不再包含任何其他更小的子博弈)的纳什均衡，并计算出相应的收益。

第二步，将每一个最后子博弈的起始点变成结束点，将计算出的每一个最后子博弈在纳什均衡下的收益写在其下方，我们就获得了一个新的扩展式博弈(或新的博弈树)，称为压缩的扩展式博弈。这样经过一次压缩，就剔除了最后子博弈。

第三步，重复第一步和第二步，并重新得到一个压缩式博弈和相应的纳什均衡。这个过程一直进行到最后只剩下惟一一个子博弈为止，这时根据在逆推过程中找到的一系列子博弈的纳什均衡组合(表明在不同信息集下的行动指南)就是扩展式博弈的一个完美均衡。

第四步，如果在逆推过程中没有遇到多重均衡，那么这个策略组合就是惟一的完美均衡；如果遇到了多重均衡，就需要对子博弈中每一个可能的均衡重复以上步骤，从而得出所有的完美均衡。比如，一个扩展式博弈有两个子博弈分别存在 2 个和 3 个纳什均衡，其他子博弈只有 1 个纳什均衡，那么该博弈就有 2×3=6 个完美均衡。

现在我们考虑一个假想的动态博弈，我们把它称为小蜈蚣博弈。

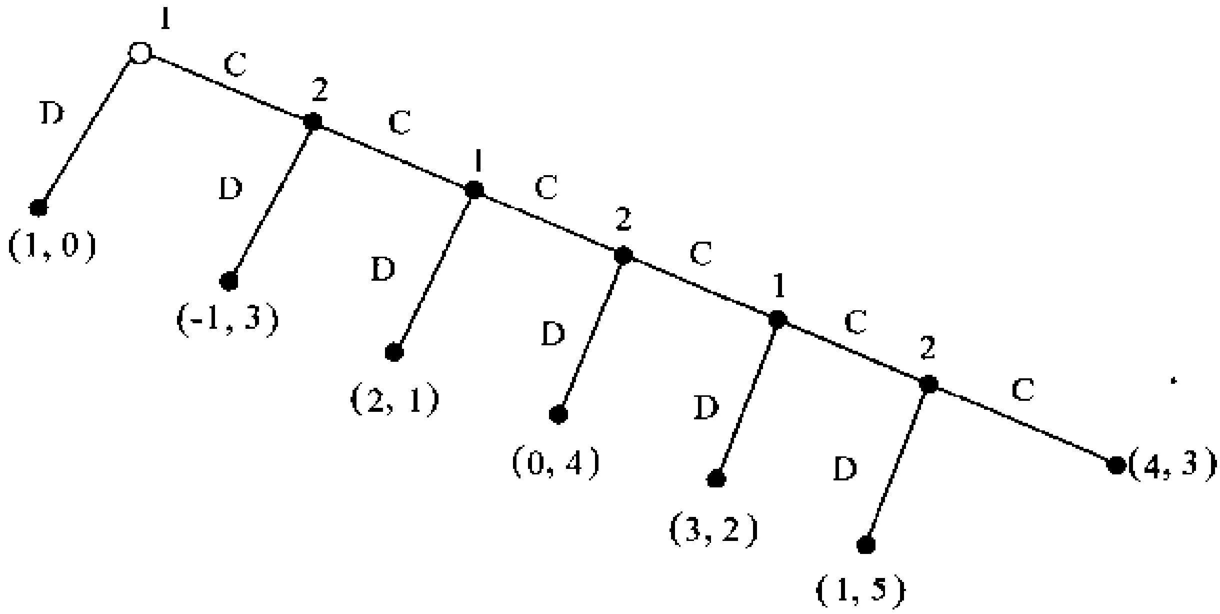


图 4.9 小蜈蚣博弈(1)

假设有两个参与者 1 和 2，他们的可用策略和收益如图 4.9 所示。每一个正在做决策的参与者，都可以决定是继续进行游戏(C)还是终止游戏(D)。参与者 1 首先决定是停止游戏，还是继续进行游戏。假如参与者 1 一开始就决定停止，则他的收益为 1，而对手的收益为 0。如参与者 1 决定继续游戏，则轮到参与者 2 作出决定。如果他选择停止，他的收益为 3，参与者 1 的收益为－1。如果他选择继续，则又轮到参与者 1 作出决定，如此等等。显然小蜈蚣博弈包含许多子博弈，如果从正面来求解小蜈蚣博弈的完美均衡会非常麻烦，特别是小蜈蚣博弈扩展到 200 次、300 次更是如此，但是用逆推法则非常简单。

我们首先求最后子博弈的纳什均衡。在最后子博弈，参与者 2 决定是选择 C 还是选择 D。如果选择 C，参与者 2 的收益为 3；如果选择 D，参与者 2 的收益为 5，所以参与者 2 在倒数第一个子博弈的纳什均衡

为 D。参与者 1 的收益为 1。现在我们将两人的收益代入倒数第二个子博弈，如图 4.10 所示。

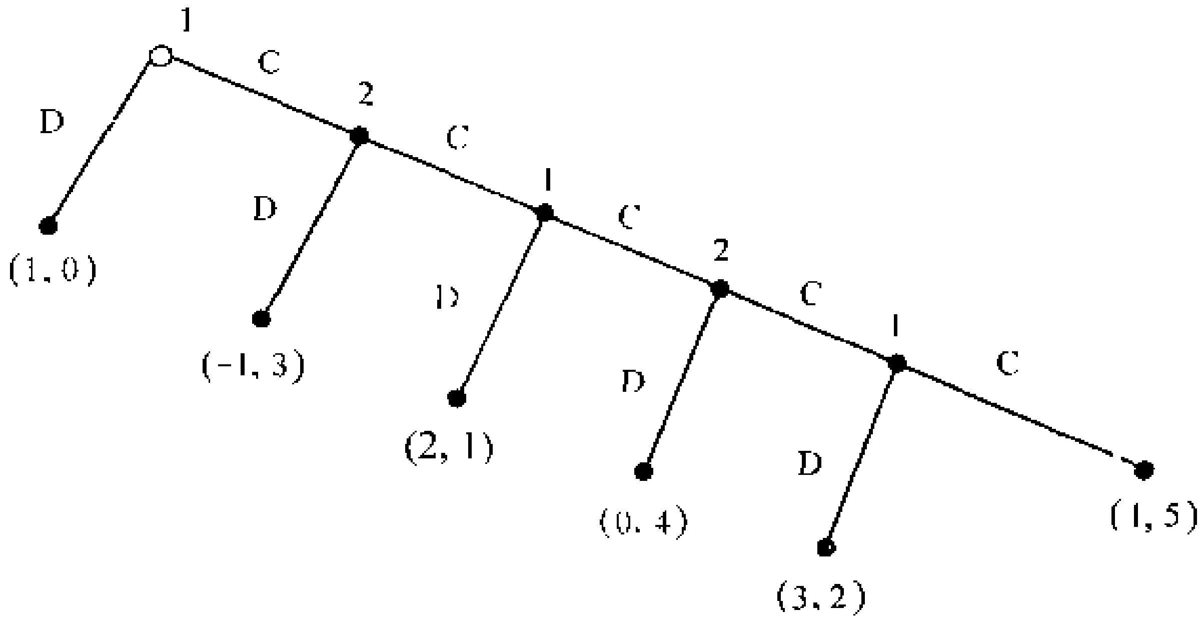


图 4.10　小蜈蚣博弈(2)

现在我们再考虑图 4.10 中的最后子博弈，它的纳什均衡为参与者 1 选择 D，收益向量为(3，2)。我们再将它带人倒数第三个子博弈。如此不断重复，一直到最后只剩下惟一一个子博弈为止，如图 4.11 所示。这时参与者 1 的最优选择为 D。

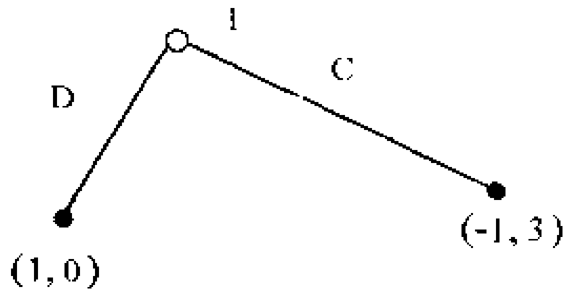


图 4.11　小蜈蚣博弈(3)

这样不断地往回解，我们可以知道任何参与者在面临选择时，都会选择结束游戏 D。因而小蜈蚣博弈的惟一的完美均衡为：参与者 1 一开始就选择结束游戏 D，如果不是参与者 2 将选择结束 D，如果不是参与者 1 将选择结束 D，如此等等，即每一个参与者都选择 D，各自的收益

为(1，0)。

实际上逆推法只不过是数学中的动态规划在扩展式博弈中的运用而已，根据动态规划理论，只要动态的阶段是有限的，最优规划(均衡)就必然存在。

定理 4.1 只要扩展式博弈 $\Gamma=\{I,X,A,p(\cdot),k(\cdot),a(\cdot),H,\rho,u\}$ 是有限的，即参与者有限，行动空间有限，博弈的阶段有限(不是无穷进行下去)，那么扩展式博弈 Γ 至少存在一个子博弈完美纳什均衡。

由于子博弈完美均衡是从最后一期的子博弈开始的，而最后子博弈就好比是一个静态博弈，又由于参与者和可用行动是有限的，因而子博弈必然存在纳什均衡(纳什定理)；然后压缩博弈，由于博弈的阶段有限，所以最后总能把完全信息扩展式博弈压缩成一个静态博弈，这就意味着扩展式博弈存在着子博弈完美均衡。如果用数学的语言来讲，由于每一个决策点的可用行动有限，因此它的概率分布(阶段策略)是一个紧致集合，在阶段策略为行为策略的向量(即混合策略)时，参与者的收益函数是连续的。由于连续函数在紧致集合上一定有极大值，[①] 所以存在一个策略使得参与者的收益最大。同样，往回走的任何一期，任何一个参与者作决策时，都可以找到最优策略。因此完美均衡存在。

接下来我们通过一些经典的例子来加深对以上内容的理解。

第三节 理论运用

在这一节里，我们将运用前面所学的理论来具体分析一些模型。

① 在数学上它被称为存在定理。

一、价格领先制

在微观经济学的垄断理论中一个有名的模型就是价格领先制，价格领先制是垄断厂商非公开勾结的普遍形式。通常领先厂商决定一个价格，然后其他厂商跟随这一价格，并选择产量。这里我们假设市场存在两个垄断企业，领先厂商（Leader）和跟随厂商（Follower）。[①] 市场的需求函数为 $P(Q)$，其中 Q 为市场总需求，q_L 为领先厂商的产量，而 q_F 则为跟随厂商的产量，它们之间的关系为 $q_L=Q-q_F$。这个博弈的时间顺序如下：(1)领先厂商选择价格 $P\geqslant 0$；(2)跟随厂商选择产量；(3)垄断企业的收益为各自的利润函数。我们不妨假设厂商的成本函数为 $C(q)$，并且 $C'(q)>0$，$C''(q)>0$。

为解出这个动态博弈的完美均衡，我们运用逆推法，首先计算跟随厂商在给定 P 下的最优产量 $q_L(P)$ 选择，即：

$$\max_{q_F} P' q_F - C_F(q_F)$$

一阶条件为：

$$P^* = C_F'(q_F) \tag{4.1}$$

即价格必须等于边际成本。由于成本函数 C' 为一一对应的连续函数，所以存在反函数，即 $q_F(P^*)=\mathrm{MC_F}^{-1}(P^*)$，$\mathrm{MC_F}$ 为边际成本。

由于领先厂商也能够像跟随厂商一样解出跟随厂商的最优反应，领先厂商就可以预测到他如选择 P^*，跟随厂商将根据 $\mathrm{MC_F}^{-1}(P^*)$来选择产量 q_F。那么，在博弈的第一阶段，领先厂商的问题就可表述为：

$$\max_{P} P[Q(P)-q_F(P)]-C_L[Q(P)-q_F(P)]$$

解一阶条件得：

$$P=C_L'(q_L)-\frac{q_L}{Q'-q_F'} \tag{4.2}$$

---

① 假设有 n 个跟随厂商和一个厂商并没有本质区别，因为他们都是价格的接收者。

由于 $Q'<0$(需求定理),$q_F'>0$(供给定理),所以上式一定大于 0。根据导数公式$\frac{\mathrm{d}f^{-1}(x)}{\mathrm{d}x}=\frac{1}{f'(x)}$,上式又等于:

$$P=C_L'(q_L)-\frac{q_L}{Q'-\frac{1}{C_F''(q_F)}} \tag{4.2.1}$$

当 q_F 越接近边际成本的最低点(回忆边际成本曲线为 U 型,停止营业点为可变平均成本的最低点,它最接近边际成本最低点),$C_F''(q_F)\to 0$,从而 $P=C_L'(q_L)$(假定 $C_L<C_F$),即领先厂商的定价如果越接近跟随厂商的停止营业点,跟随厂商的产量就越小,从而领先厂商的市场份额就越大,当领先厂商的定价等于跟随厂商的停止营业点时,那么跟随厂商的产量为零,而领先厂商占有整个市场,所以他的定价等于边际成本,即 $P=C_L'(q_L)$。由于我们假定模型为完全信息,所以只要给定需求函数和成本函数,我们就可以得到具体的值。因而((4.1),(4.2))就是价格领先制的完美均衡。

如果我们把上述模型稍加改造——领先厂商不是选择价格而是选择产量,我们会得出一个有趣的现象,即领先厂商占有更多的信息反而带来了不利,而这种情况在静态博弈中不会出现。在课后习题中读者将按要求在产量领先制中证明这个观点。

二、工资和就业博弈

在一个劳动力市场上存在一个垄断性质的工会,工会决定工资率,而企业根据工资来决定需要雇用多少工人。工会的效用函数为 $U(w,L)$,其中 w 为工会向企业开出的工资水平,L 为就业量。假定 $U(w,L)$都是 w 和 L 的增函数,即工资率越高,就业量越高,工会取得的成绩越大。企业的利润函数为 $\pi(w,L)=R(L)-wL$,其中 $R(L)$为企业雇用 L 名工人在最优的生产和产品市场决策下可获得的收入,假定 $R(L)$为凹函数,即 $R'(L)>0$,$R''(L)<0$。如图 4.12 所示。

该博弈的时序如下:(1)工会决定劳动力市场的工资率;(2)企业在工资率 w 下,决定雇用的工人人数;(3)工会和企业各自得到自己的收

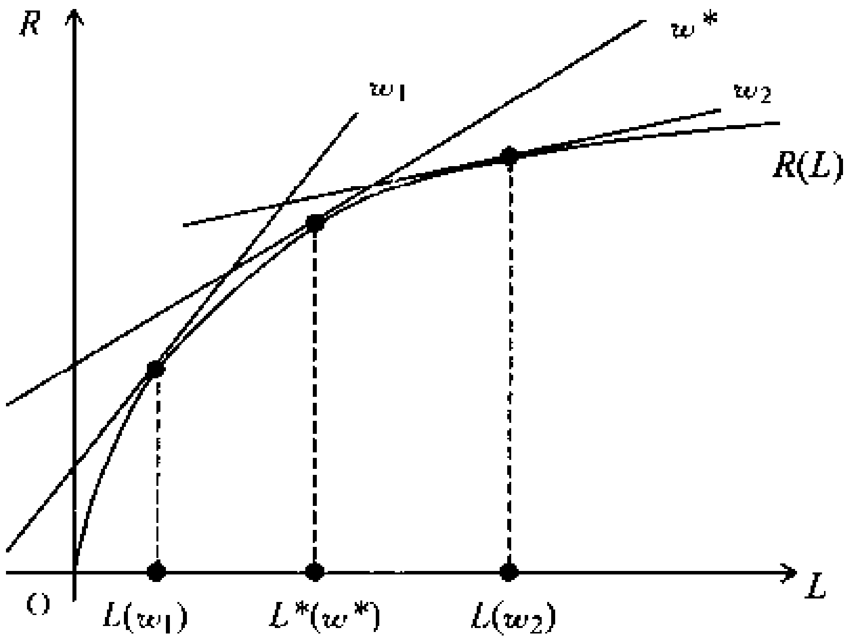


图 4.12 凹函数

益，收益分别为 $U(w,L)$ 和 $\pi(w,L)$。利用逆推法解该动态博弈的完美均衡。

首先，在给定工会在第一阶段决定的工资率 w 的条件下，企业应选择 $L(w)$ 使利润最大化：

$$\max_{L\geq 0} R(L)-wL$$

一阶条件为：

$$R'(L^*)-w^*=0 \tag{4.3}$$

由于 $R(L)$ 为凹函数，上式必有解，并且一阶条件也是充分条件。求(4.3)式的反函数，我们就能得到 $L(w)$，它实际上表示在利润最大化下，L 如何根据工资率 w 来选择，即给定 w，那么必然存在 L 使得利润最大。根据(4.3)式和 R 为凹函数，我们可知，当 $L>L^*$ 时，$R'(L)-w^*<0$；而当 $L<L^*$ 时，$R'(L)-w^*>0$，如图 4.12 所示。这表明企业的等利润曲线是一个倒 U 型，$L(w)$ 必定过各等利润曲线的最高点。由此可画成图 4.13。

当 L 保持不变，工资率 w 降低会使企业的利润提高，所以较低的等利润曲线代表较高的利润水平，在图 4.13 中我们用箭头标出。

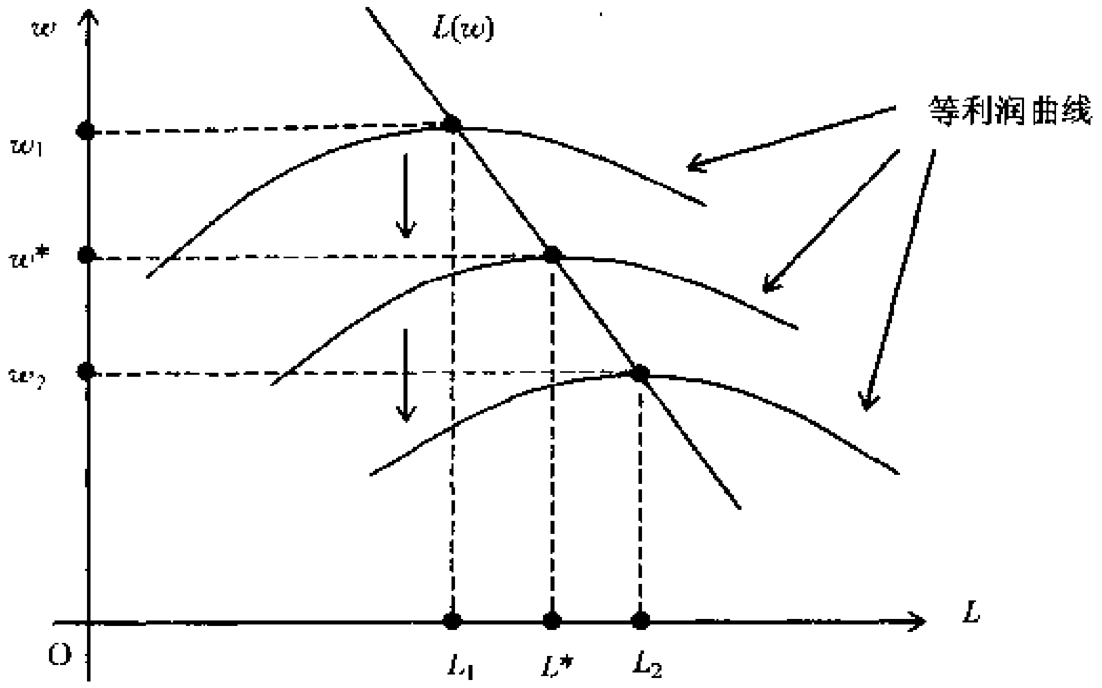


图 4.13 企业的等利润曲线

最后我们分析工会在第一阶段的问题。工会在第一阶段的问题是根据企业的决策 L,选择工资率以使工会的效用函数最大化,即:

$$\max_{w\geq 0} U(w, L^*(w))$$

由于我们不知道效用函数的具体形式,因而我们用工会的无差异曲线来代替上述公式最大化。根据微观经济学的知识,当无差异曲线与约束线相切时,切点所代表的组合将使效用最大化。由上面的分析可知,图 4.13 中的 $L(w)$ 就是工会面临的约束线,因为企业追求利润最大化必然导致这一结果。在图 4.14 中,工会的最优选择就是使它的无差异曲线与 $L(w)$ 曲线相切,(w^*, L^*) 就为这一工资和就业博弈的完美均衡。

但这个完美均衡是无效率的,如果 w 和 L 位于阴影部分之内,则工会和企业的福利都能得到改善。上述分析的结论成为批评工会控制工资率导致无效率的一个重要理由。在后面介绍无名氏定理的时候,我们将看到如果博弈无限期地重复进行(或不知道博弈什么时候结束),那么完美均衡就有可能在阴影部分之内出现。

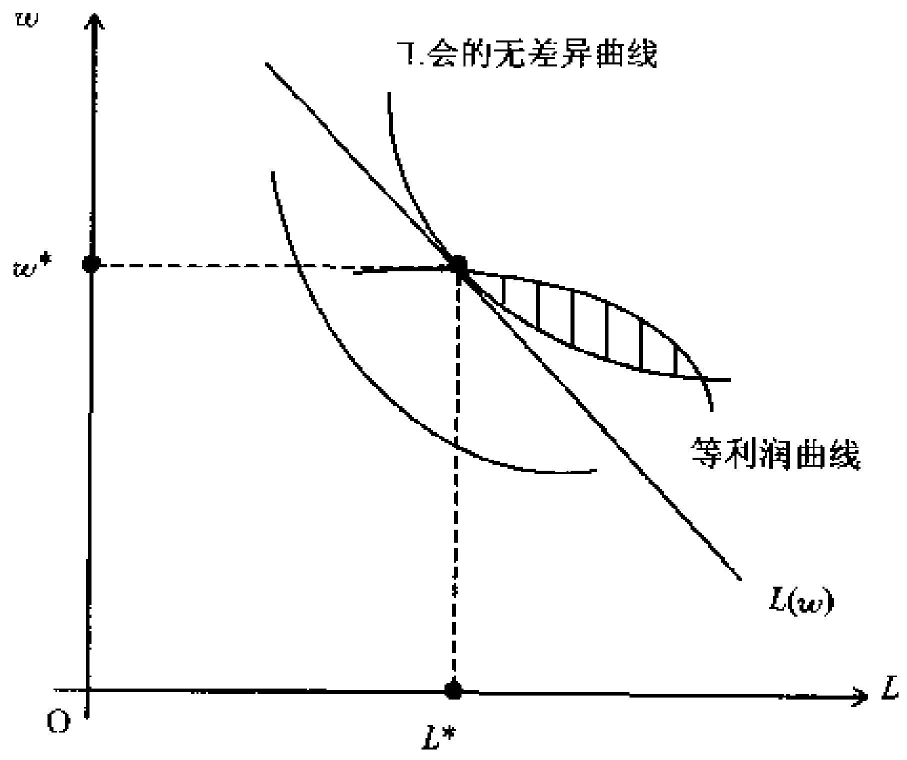


图 4.14　最优选择与无效率

三、搜寻模型

搜寻模型在宏观经济学中得到了广泛的运用，为分析人们在找工作过程中是如何失业的提供了一种理论解释。这里我们假设有一位大学毕业生，他开始寻找工作，我们不妨假设有 4 个公司通知他面试。我们假设公司的工资报价服从$[0,L]$的均匀分布，L 是公司可能给出的最高工资。为了简化分析，我们这里假设公司在支付工资不超过 L 的情况下总是愿意接受新工人，并且省掉了公司的收益。博弈的程序如下：第一阶段，第 1 个公司的行动是给出工资报价，而毕业生决定是否接受该工资。如果接受，博弈结束；如果不接受，毕业生到下一个公司面试。第二阶段，第 2 个公司的行动也是给出工资报价，而毕业生的策略是接受或不接受，如接受博弈结束，如不接受毕业生转入下一阶段的面试，如此等等直到第 4 个公司。我们不妨假设毕业生有一个保留工资 U，如果最后获得的工资低于保留工资，那么毕业生就可能选择失业（例如考研究生）而不是工作。由于每一个公司工资报价的期望值都等

于 $L/2$，因而毕业生的保留工资 U 为工资报价的期望值，即 $U=L/2$。

我们运用逆推法，从第 4 个公司开始。如果毕业生在第 1 个公司、第 2 个公司、第 3 个公司都没有接受工资报价，那么在第 4 个公司，它能获得的期望值为：

$$\mathrm{E}w = \int_0^L w f(w)\mathrm{d}w = \int_0^L w \frac{1}{L}\mathrm{d}w = \frac{L}{2}$$

对于毕业生而言，如果第 4 个公司的工资报价大于等于 $L/2$，那么他接受工作，否则选择失业。即在第 4 阶段，毕业生的最优策略为工资报价低于 $L/2$ 拒绝工作，工资报价高于等于 $L/2$ 接受工作。第 4 个公司的最优策略则为 $w=L/2$。

在第三阶段，如果毕业生在第 1 个公司和第 2 个公司都没有接受工资报价，那么在第 3 个公司，它能获得的期望值为：

$$\begin{aligned}\mathrm{E}w &= \int_0^L \max[w, L/2]\,\frac{1}{L}\mathrm{d}w \\ &= \int_{L/2}^L \frac{w}{L}\mathrm{d}w + \int_0^{L/2} \frac{L}{2} \times \frac{1}{L}\mathrm{d}w = \frac{5L}{8}\end{aligned}$$

由于毕业生在第 4 个公司能获得工资的期望值为 $L/2$，所以只有第 3 个公司提供的工资高于 $L/2$，毕业生才会接受。在第三阶段纳什均衡为：如果第 3 个公司出价大于等于 $5L/8$，那么毕业生接受工作，否则到第 4 个公司面试。而第 3 个公司的最优工资报价为 $5L/8$。

在第二阶段，如果毕业生拒绝了第 1 个公司的报价，那么在第 2 个公司，它能获得的期望值为：

$$\begin{aligned}\mathrm{E}w &= \int_0^L \max[w, 5L/8]\,\frac{1}{L}\mathrm{d}w \\ &= \int_{5L/8}^L \frac{w}{L}\mathrm{d}w + \int_0^{5L/8} \frac{5L}{8} \times \frac{1}{L}\mathrm{d}w = \frac{89L}{128}\end{aligned}$$

即由于毕业生在第 3 个公司的期望工资为 $5L/8$，只有第 2 个公司提供的工资大于等于 $5L/8$，毕业生才会接受。因而第 2 阶段的纳什均衡为：如果第 2 个公司的报价大于等于 $89L/128$，那么毕业生接受工

作，否则到第 3 个公司面试。而第 2 个公司的最优工资报价为 $89L/128$。

在第一阶段，毕业生期望能获得的工资报价为：

$$Ew=\int_0^L \max[w,89L/128]\frac{1}{L}dw$$

$$=\int_{89L/128}^{L}\frac{w}{L}dw+\int_0^{89L/128}\frac{89L}{128}\times\frac{1}{L}dw\approx 0.742L$$

即如果第 1 个公司的工资报价为 $0.742L$，那么毕业生接受工作，否则到第 2 个公司面试，而第 1 个公司的最优报价为 $0.742L$。因此我们看到，搜寻博弈的完美均衡为：如果第 1 个公司的工资为 $0.742L$，那么毕业生接受工作，否则到第 2 个公司面试；如果第 2 个公司的工资报价为 $0.695L$，那么毕业生接受工作，否则到第 3 个公司面试；如果第 3 个公司的工资报价为 $0.625L$，那么毕业生接受工作，否则到第 4 个公司面试；如果第 4 个公司的工资报价为 $0.5L$，那么毕业生接受工作，否则选择失业。这里的一个有趣现象是，如果毕业生的选择机会越多，那么通常对工资的预期值越高，因而失业的可能性也就越大。这就从微观层面上为宏观经济学中的失业提供了一种说明，而且我们看到失业可以作为完美均衡的一种结果出现。

上面我们介绍的动态博弈都是完美信息下的博弈，即后来者观察到所有前行者的行动，也即所有信息集都是单点信息集。对于完美信息博弈我们有如下定理：

定理 4.2　如果扩展式博弈 Γ 为完美信息博弈，那么它存在惟一的子博弈完美纳什均衡。如果在扩展式博弈 Γ 中，任何一个参与者在各结束点的收益都不相等，那么存在惟一的纯策略组成的子博弈完美纳什均衡。

下面我们考察稍微复杂一点的模型——非完美但完全信息动态博弈。在博弈的某一阶段，有两个参与者同时行动，所以这是一个非完美但完全信息的博弈（我们曾说过所有完全信息静态博弈都是非完美信息博弈）。

四、银行挤兑(Bank run)

在现实生活中，不时发生的银行危机的一个重要原因就是发生了大规模的银行挤兑，从而导致银行发生连锁性倒闭，造成严重的金融危机。其典型的例子就是1929年经济大萧条中的银行倒闭。银行的一个重要功能就是吸收短期零散的存款，并把它作为长期贷款进行投资，这就形成了银行的一个重要特征——部分储备。银行的部分储备制不仅是中央银行实行宏观货币策略的基础，也为各种银行危机的产生客观上提供了技术条件。一旦发生挤兑，(如果没有中央银行的救援)通常银行只能倒闭。

这里我们分析一个简单的银行挤兑模型。假设有两个存款人每人存入银行一笔存款D，银行将这些存款投入一个长期项目，只有在第二时期，这个项目才能够完成。为了分析简化，我们假设没有贴现。在第一个时期两个存款人可以提款，也可以不提款，如果提款则意味着银行被迫收回投资，能收回的钱为2r，这里$D>r>D/2$。如果只有一个人在第一时期提款，那么他得到D，另一个人得到$2r-D$。如果两个人都提款，各自得到r。如果两人都不提款，那么进入第二时期。在第二个时期，如果两人都提款，那么每人得R，其中$R>D$(连本带利)。如果只有一人提款，那么提款人得$2R-D$，不提款人得D。如果两人都不提款，则银行向每个存款者返还R。在第二期时博弈结束。这个博弈的扩展式表示如图4.15。其中W表示提款，S表示不提款。

这个模型存在两个子博弈，利用逆推法，从最后一个子博弈开始。在第二期，如果存款人1选择提款，那么存款人2的最优选择是提款；如果存款人1选择不提款，那么存款人2的最优选择仍是提款，即提款是存款人2的严格优策略。对于存款人1而言，提款也是严格优策略。因而最后一个子博弈的纳什均衡为(提款，提款)，相应的收益为(R,R)。然后压缩博弈树，就得到图4.16。

从图4.16可以看出，这个子博弈存在两个纳什均衡：一个是(提款，提款)，一个是(不提款，不提款)。综上所述，在银行挤兑这个模型中，存在着两个子博弈完美纳什均衡解：(1)在第一期和第二期，两个存款人都

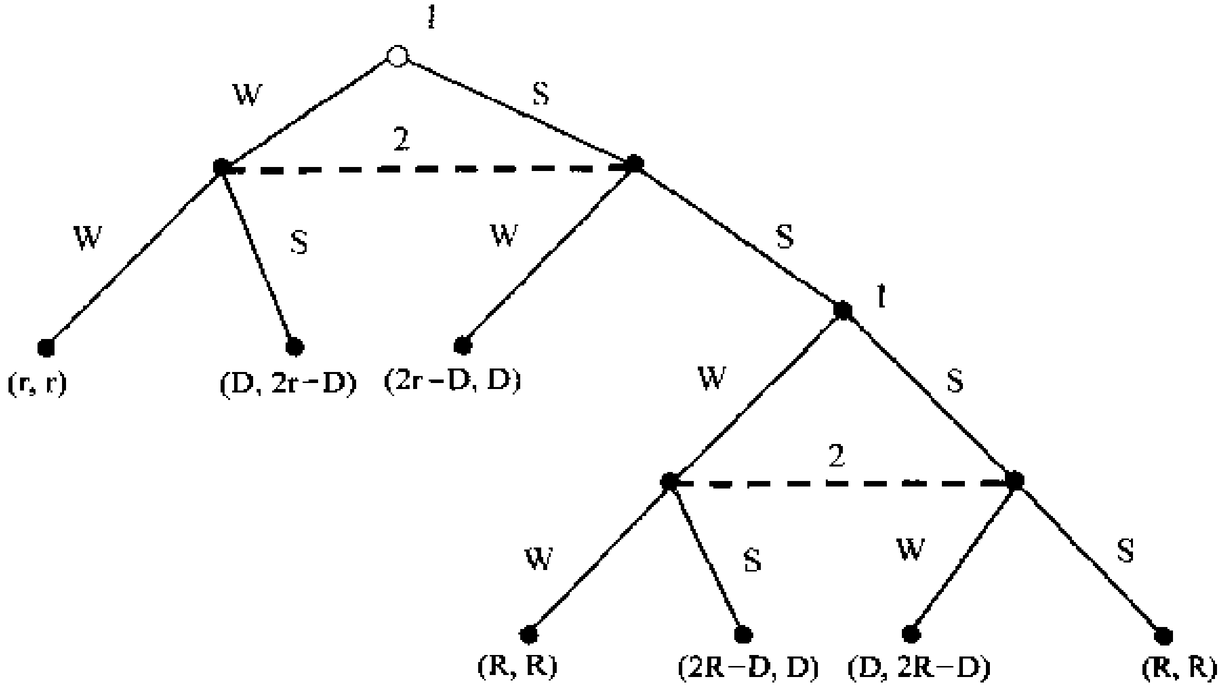


图 4.15　银行挤兑(1)

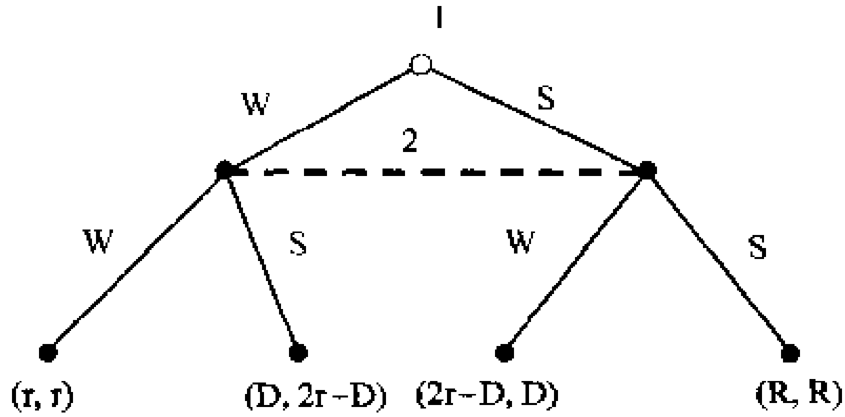


图 4.16　银行挤兑(2)

提款，即{(W，W)，(W，W)}，收益分别为(r，r)。实际上由于第一期提款后，就到不了第二期，但我们仍需要说出参与者的纳什均衡。(2)在第一期，两个存款人都不提款，而在第二期都提款，即{(S，W)，(S，W)}，收益分别为(R，R)。第一个完美均衡实际上就是对银行的一次挤兑。由于还存在着另一个均衡，因而这个模型并不能预测发生银行挤兑的概率，一个解决的办法就是在模型中引入信息不对称，假设存款人对银行的挤兑依赖于对银行投资质量的判断，它服从一定的概率分布，银行的投资

质量低，投资失败的可能性就大，从而银行挤兑和金融危机的发生就服从一定的概率分布。20 世纪 90 年代大量的有关金融危机的文献就是从以上这个方面来建立银行危机模型的。由于银行挤兑模型是一个非完美信息博弈，因而出现了两个完美均衡，这也间接地证明了定理 4.2。

五、国际贸易与关税

各国国家之间的贸易政策和关税协定实际上是各国间博弈的结果，中国加入世贸组织的谈判过程就清楚地表明了这一点。这里我们考虑一个简单的关税博弈模型。假设有两个完全相同的国家，定义为 $i=1,2$。每个国家有一个政府来制定关税税率，一个企业为国内市场和国外市场生产一种产品。如果 $i(=1,2)$国的市场总产量为 Q_i，则市场需求为 $p_i(Q_i)=a-Q_i$，国家 i 中的企业为国内市场生产 h_i，出口 e_i，则 $Q_i=h_i+e_j$，即 i 国的总需求量为国内产量加进口。企业的边际成本为常数 c，没有固定成本，那么企业 i 的成本函数为 $C_i(h_i,e_i)=c(h_i+e_i)$。另外，产品出口时还要负担关税成本：如果政府 j 制定的关税为 t_j，那么企业 i 向 j 国出口 e_i 必须支付的关税为 t_je_i。

该博弈的时间顺序如下：(1)政府同时选择关税税率 t_1 和 t_2；(2)两国的企业 1 和 2 观察到关税，并确定产量 h_i 和 e_i；(3)企业的收益为利润最大化，而两国政府的收益为福利最大化，福利由消费者剩余、企业利润和关税组成。根据微观经济学的知识，消费者剩余为消费者获得的总效用减去支出。由于需求曲线是线性的，消费者剩余为：

$$\int_0^Q[(a-x)-p]\mathrm{d}x=\int_0^Q[(a-x)-(a-Q)]\mathrm{d}x$$

$$=-\frac{1}{2}Q^2+Q^2=\frac{1}{2}Q^2$$

$$=\frac{1}{2}(h_i+e_j)^2$$

第 i 国企业的利润函数为：

$\pi_i(t_i,t_j,h_i,e_i,h_j,e_j)$

$$=[a-(h_i+e_j)]h_i+[a-(e_i+h_j)]e_i-c(h_i+e_i)-t_je_i$$

第 i 国政府的福利函数为：

$$w_i(t_i,t_j,h_i,e_i,h_j,e_j)=\frac{1}{2}Q_i^2+\pi_i(t_i,t_j,h_i,e_i,h_j,e_j)+t_ie_j$$

根据逆推法，我们首先考虑在给定两国政府制定的关税税率 t_1 和 t_2 的条件下，两个企业的纳什均衡。对于每一个企业 i，(h_i,e_i) 必须满足：

$$\max_{h_i,e_i} \pi_i(t_i,t_j,h_i,e_i,h_j,e_j)$$

由于企业的利润是两个市场利润之和，企业 i 在市场 i 的利润只是 h_i 和 e_j^* 的函数，而在市场 j 的利润却是 h_j^*，e_i 和 t_j 的函数，企业 i 的利润最大化函数可以拆分为两个问题，即在每一个市场分别求最大化。

国内市场 h_i 必须满足：

$$\max_{h_i} h_i[a-(h_i+e_j^*)-c]$$

国际市场 e_i 必须满足：

$$\max_{e_i} e_i[a-(h_j+e_i)-c]-t_je_i$$

假设 $e_j^*\leqslant a-c$ 和 $h_j^*\leqslant a-c-t_j$，可得：

$$\begin{cases} h_i^*=\frac{1}{2}(a-e_j^*-c) & (4.4)\\ e_i^*=\frac{1}{2}(a-h_j^*-c-t_j) & (4.5)\end{cases}$$

对于每一个企业 $i=1,2$，都必须满足(4.4)式和(4.5)式两个最优反应函数，从而对四个未知数 (h_1,e_1,h_2,e_2) 存在四个方程式，求解非常容易。其解为：

$$h_i^*=\frac{a-c+t_i}{3} \quad 和 \quad e_i^*=\frac{a-c-2t_j}{3} \qquad (4.6)$$

从上述两个解可以看出，如果没有关税，那么两个企业生产的产量等于古诺模型中的产量，即在国内和国外都生产 1/3 的市场容量。但由

于关税的存在，使得国内企业的生产多增加了 $t_i/3$ 的产量，即受到了国内关税的保护。相反，在国外由于关税保护而受到了损失，产量减少 $2t_j/3$。此外，h_i^* 随着国内关税 t_i 的提高而增加，进口 e_j^* 则随着国内关税的提高而减少，因而关税的提高实际上会增加国家间的封闭性，而不会是外向性，这实际上是一个囚徒困境。

在第一阶段，政府要解出最优的税率使福利最大化。由于 $(h_1^*, e_1^*, h_2^*, e_2^*)$ 都是税率的函数，所以在福利函数中它们被省略掉了。

$$\max_{t_i \geq 0} w_i(t_i, t_j)$$

$$\max_{t_i} \frac{[2(a-c) \quad t_i]^2}{18} + \frac{(a-c+t_i)^2}{9} + \frac{(a-c-2t_j^*)^2}{9} + \frac{t_i(a-c-2t_i)}{3}$$

求一阶条件，得：

$$t_i^* = \frac{a-c}{3}$$

这一结果对每一个 i 都成立，并不依赖于 t_j，这表明各国政府选择 1/3 的税率是严格优策略。把 $t_1 = t_2 = (a-c)/3$ 带入(4.6)式可得：

$$h_i^* = \frac{4(a-c)}{9} \quad 和 \quad e_i^* = \frac{a-c}{9}$$

这就得到企业在第二阶段所选择的产量。因而，关税博弈中的子博弈完美纳什均衡解为：第一阶段各国政府选择 $(a-c)/3$ 的关税，第二阶段企业在国内生产 $4(a-c)/9$，在国外生产 $(a-c)/9$，即 $[t_1^* = t_2^* = (a-c)/3, h_1^* = h_2^* = 4(a-c)/9, e_1^* = e_2^* = (a-c)/9]$。

现在我们不妨先假设关税为零，那么各国市场上的商品总量为 $2(a-c)/3$，消费者剩余为 $4(a-c)^2/9 = 32(a-c)^2/81$，而在税率为 $(a-c)/3$ 的情况下，各国的商品总量为 $5(a-c)/9$，消费者剩余为 $25(a-c)^2/81$，即零关税带来更大的消费者剩余。其次，我们求各国福利之和的最大值，即：

$$\max_{t_1,t_2} w_1(t_1,t_2)+w_2(t_1,t_2)$$

利用我们前面分析的结果，解一阶条件，得：

$$-(a-c)=t_i$$

由于税率不可能为负数，所以 $t_1=t_2=0$，即实行零关税，不仅能增加各国的福利，而且能够增加两国的总体福利。由此我们看到，实行关税是一个完美均衡，但对整个社会而言却是低效率的。于是，各国政府就有充分的理由来签订一个相互承诺零关税的协定，这就是为什么WTO能够不断发展的根本原因（如果负关税税率代表出口补贴，社会最优的条件是政府选择 $t_1=t_2=-(a-c)$，这使得国内企业为本国消费者提供的产出为零，并向另一国家出口完全竞争条件下的产量 $a-c$）。

六、工作竞赛

假设有两个工人，每一个工人 i 的产出取决于努力程度 e_i 和随机扰动项 ε_i，即 $y_i=e_i+\varepsilon_i$。生产的程序如下：第一，两个工人同时选择非负的努力水平 $e_i\geqslant 0$；第二，随机扰动项 ε_1 和 ε_2 相互独立，服从期望值为0，密度函数为 $f(\varepsilon)$ 的概率分布；第三，工人的产出可以观察到，但工人的努力水平无法观察，从而工人的工资只能决定于个人的产出，而无法直接取决于其努力水平。

假设企业主为了激励工人努力工作，在他们之间开展工作竞赛。产出高的工人即是工作竞赛的优胜者，获得高工资 w_H，而失败者获得低工资 w_L。工人获得工资水平 w 并付出努力程度 e 时的收益为 $u(w,e)=w-g(e)$，其中 $g(e)$ 表示努力工作带来的负效用，是递增的凸函数，即 $g'(e)>0$，且 $g''(e)>0$。企业主的收益为：$y_1+y_2-w_H-w_L$。

我们定义企业主为参与者1，他的行动是选择工作竞赛中的工资水平 w_H 和 w_L，两个工人为参与者2和参与者3，他们观察到第一阶段选定的工资水平，然后同时选择努力程度 e_1 和 e_2。最后，参与者各自的收益如前面所给出的。由于产出不仅是参与者行动的函数，而且还受到随机扰动因素 ε 的影响，因而参与者的收益为期望值。

我们仍运用逆推法。给定企业主选定的工资率 w_H 和 w_L，如果一

对努力水平(e_1^*,e_2^*)为第二阶段的纳什均衡，则对每一工人 i 而言，e_i 一定是使他的期望工资减去努力带来的负效用后的净收益最大化的解，即：

$$\max_{e_i} w_H\text{Prob}\{y_i(e_i)>y_j(e_j)\}+w_L\text{Prob}\{y_i(e_i)\leqslant y_j(e_j)\}-g_i(e_i)$$
$$=(w_H-w_L)\text{Prob}\{y_i(e_i)>y_j(e_j)\}+w_L-g(e_i) \tag{4.7}$$

其中 $y_i=e_i+\varepsilon_i$。

(4.7)式的一阶条件为：

$$(w_H-w_L)\frac{\text{dProb}\{y_i(e_i)>y_j(e_j^*)\}}{\text{d}e_i}=g'(e_i) \tag{4.8}$$

(4.8)式实际上就是边际收益必须等于边际成本，边际收益为奖励工资(w_H-w_L)与获得奖励工资的概率的乘积(期望值)，而边际成本就是边际负效用。

根据贝叶斯法则(记住独立性)，则：

$$\begin{aligned}\text{Prob}\{y_i(e_i)>y_j(e_j)\}&=\text{Prob}\{\varepsilon_i>e_j+\varepsilon_j-e_i\}\\&=\int_{-\infty}^{+\infty}\text{Prob}\{\varepsilon_i>e_j^*+\varepsilon_j-e_i|\varepsilon_j\}f(\varepsilon_j)\text{d}\varepsilon_j\\&=\int_{-\infty}^{+\infty}[1-F(e_j^*+\varepsilon_j-e_i)]f(\varepsilon_j)\text{d}\varepsilon_j\end{aligned}$$

于是(4.8)式可化为：

$$(w_H-w_L)\int_{-\infty}^{+\infty}f(e_j^*+\varepsilon_j-e_i)f(\varepsilon_j)\text{d}\varepsilon_j=g'(e_i)$$

由于我们不清楚具体的函数形式，因而这里要运用到一个技巧，即由于随机扰动项对两个工人的影响的期望值是一样的，因而如果给定工人 j 的努力程度为 e_j，那么工人 i 的努力程度 e_i 只要比 e_j 大一个无穷小正数就能获胜(从期望值来看)，因而最优的努力程度是 $e_i=e_j$，所以第二阶段的纳什均衡意味着 $e_i^*=e_j^*$。从而上式可化为：

$$(w_H-w_L)\int_{-\infty}^{+\infty}f(\varepsilon_j)^2\text{d}\varepsilon_j=g'(e_i^*) \tag{4.9}$$

由于 $g(e)$ 是凸函数，优胜者获得的奖励越高(即 $w_H - w_L$)，就越会激励工人更大的努力，这和我们的直觉是一致的。另一方面，在同样的奖励水平下，若产生的随机扰动因素越大，就越不值得努力工作，因为这时工作竞赛的最终结果很大程度上取决于运气而不是努力。例如，当 ε 服从正态分布时，则有：

$$\int_{-\infty}^{\infty} f(\varepsilon_j)^2 d\varepsilon_j = \frac{1}{2\sigma\sqrt{\pi}}$$

它随 σ 的增加而下降，也就是说 e_i^* 随 σ 的增加面下降(根据(4.9)式)。

在给定工人的最优选择下，企业主如何决定 w_H 和 w_L？如果工人同意参与者工作竞赛而不是另谋高就，他们对给定的 w_H 和 w_L 就会是(4.9)式所给出的纳什均衡策略。同时，假定工人另谋高就能得到的保留效用为 U。因为在对称的纳什均衡中($e_i^* = e_j^*$)，每个工人在竞赛中获胜的概率为 $\text{Prob}\{y_i > y_j\} = 1/2$，如果企业主要使工人有积极性参加工作竞赛，则它选择的工资率必须满足下式：

$$\frac{1}{2} w_H + \frac{1}{2} w_L - g(e^*) \geq U \tag{4.10}$$

(4.10)式在经济学中被称为参与约束。它表示要吸引人参与博弈必须满足的条件。假如 U 足够低，以至于工人愿意参加竞赛，则企业主在(4.10)式的约束下，选择使自己期望收益 $2e^* - w_H - w_L$ 最大的工资水平(随机扰动项的期望值为 0)。由于在最优条件下，(4.10)式中的等号成立：

$$w_L = 2U + 2g(e^*) - w_H \tag{4.11}$$

则期望利润就成为 $2e^* - 2U - 2g(e^*)$，于是企业主要考虑的问题就是使 $e^* - g(e^*)$ 最大化，这时他选择的工资水平应激励工人选择 e^* 满足这一条件。从而最优条件下的努力程度满足一阶条件 $g'(e^*) = 1$，将其代入(4.9)式则意味着最优激励 $w_H - w_L$ 满足：

$$(w_H - w_L)\int_{-\infty}^{+\infty} f(\epsilon_j)^2 \mathrm{d}\epsilon_j = 1$$

即 $w_H - w_L = 2\sigma\sqrt{\pi}$。连同(4.11)式一起，可解出 w_H 和 w_L 的值。因而该博弈的完美均衡解为：(w_H^*, w_L^*)，$e_1^* = e_2^*$，其中 $g'(e^*) = 1$。

第四节　重复博弈

在博弈论中，如果博弈是无穷进行的(或不知道博弈何时结束)，那么我们根本无法判断是否存在纳什均衡，这就是为什么我们无从得知象棋的均衡解，如果能够得知，国际象棋、围棋的比赛也就没有什么意义了，尽管这些比赛是完美信息博弈。但有一种博弈例外，即重复博弈。对重复博弈的分析使我们对人类社会普遍存在着的各种利己利人的合作制度有了更为深刻的理解。

一、有限重复博弈

顾名思义，重复博弈就是相同的博弈 G 在长期中不断地重复进行，动态博弈的任何阶段实际上都是一个完全信息博弈。我们把动态博弈中具有相同片段的博弈就称为阶段博弈。正式的定义如下：

定义 4.8　对于完全信息博弈 $G=\{I,S,u\}$，其中 $I=\{1,2,\cdots,n\}$ 为参与者集合，$S=\{S_1,\cdots,S_n\}$ 为所有参与者的策略空间，$u=\{u_1,\cdots,u_n\}$ 为所有参与者的收益函数，如果 G 在时间上(程序上)不断重复，并且在下一次博弈 G 开始前，所有以前博弈的历史都被观察到，那么它构成的动态博弈就称之为重复博弈，G 就为重复博弈中的阶段博弈。如果 G 重复进行 T 次，那么 G(T)就表示重复进行 T 次的有限重复博弈。如果 G 重复进行∞次，那么 G(∞)就表示无限重复博弈。

实际上对于重复博弈中的阶段博弈 G，不仅为完全信息静态博弈，也可以为完全信息动态博弈；不仅可以为完全信息博弈，也可以为非完全信息博弈。由于有限重复博弈是对 G 重复 T 次，因而分析实际上取决

于 G。如果阶段博弈 G 存在纳什均衡，那么 G(T)的子博弈完美纳什均衡不过是纳什均衡重复 T 次，根本的原因是 G(T)只不过是 G 的有限次重复。如果 G 有多个纳什均衡，G(T)的完美均衡就有许多种可能。

定理 4.3　如果阶段博弈 G 有惟一的纳什均衡，则对任意有限的 T，重复博弈 G(T)有惟一的子博弈完美纳什均衡，即 G 的纳什均衡结果在每一个阶段重复进行。

这里我们考虑一个简单的两阶段重复博弈，即重复两次的信用困境博弈，如图 4.17 所示。如果两个商人都讲诚信的话，那么他们的福利都能改善，各得 4 元人民币；如果一方讲诚信而另一方欺骗，那么欺骗方能获得更大的利益(5 元)。这个博弈与囚徒困境实质上完全一样，惟一的纳什均衡为(欺骗，欺骗)，是一个严格优策略均衡。

		商人 2 诚信	商人 2 欺骗
商人 1	诚信	4，4	0，5
	欺骗	5，0	1，1

图 4.17　信用困境(1)

根据逆推法我们知道，在第二阶段，信用困境的纳什均衡为(欺骗，欺骗)，而不管第一阶段的选择如何。逆推到第一阶段，将第二阶段的纳什均衡收益代人，则如图 4.18 所示。

		商人 2 诚信	商人 2 欺骗
商人 1	诚信	5，5	1，6
	欺骗	6，1	2，2

图 4.18　信用困境(2)

在第一阶段，惟一的纳什均衡还是(欺骗，欺骗)。从而，两阶段信用困境惟一的完美均衡为第一阶段的(欺骗，欺骗)和随后第二阶段的(欺骗，欺骗)。利己利人的合作解——(诚信，诚信)根本不可能在有限重复博弈中出现。如果我们假设商业信用博弈重复 100 次、200 次、10000 次，上述两阶段分析的结果仍然不会改变。因为根据逆推法，G 中惟一

的纳什均衡在 G(T)中将重复出现 T 次。这个结果一方面与现实中的一些情况吻合，即商业交往中存在着普遍的商业欺骗，商业欺骗已经成为中国在向市场经济转变过程中出现的一个毒瘤，但另一方面也存在冲突。因为现实中还存在着大量讲诚信的商人，(诚信，诚信)这样的合作广泛地存在于经济生活中，否则人类社会早已终结。然而，有限重复博弈理论排除了这种可能。为了在理论上容纳合作解，博弈论主要从三个方面来加以发展：一是引入多重均衡；二是引入无限重复博弈；三是引入信息不完全。20 世纪 90 年代以来博弈论的理论研究重点基本上就是按照上述三个方面展开的。

现在，我们仍回到两阶段博弈，进一步考察阶段博弈 G 存在多个纳什均衡的情况，如图 4.19 所示。这个博弈与原来信用困境不同的地方是人为地构造了一个新的纳什均衡——(中，中)，因而这个阶段博弈存在两个纳什均衡——多重均衡。

		商人 2		
		诚信	欺骗	中
商人 1	诚信	4,4	0,5	0,0
	欺骗	5,0	1,1	0,0
	中	0,0	0,0	3,3

图 4.19 信用困境(3)

我们仍然规定图 4.19 所表示的阶段博弈重复进行两次。与原来版本一个重要的不同是合作解有可能成为子博弈完美解在重复博弈中出现。现在我们就证明在信用博弈 1 这个重复博弈中存在一个子博弈完美纳什均衡解：其中第一阶段的策略组合为(诚信，诚信)，第二阶段的策略组合为如果对方在第一阶段选择了诚信，那么选择中，如果没有选择诚信，那么选择欺骗。根据逆推法，我们首先考察第二阶段。

在第二阶段，(欺骗，欺骗)和(中，中)都是纳什均衡，因面策略组合[(如果对手诚信，选择中；如果非诚信，选择欺骗)，(如果对手诚信，选择中；如果非诚信，选择欺骗)]是阶段博弈 G 的一个纳什均衡。把这一均衡对应下的收益代入第一阶段，就得到图 4.20。图 4.20 中的收益是

这样得到的，第一阶段的策略组合(诚信，诚信)对应的收益要各加上3(因为在第一阶段选择了诚信，得到奖励)，而其余的8个单元各加上1(因为选择了非诚信，受到惩罚)。实际上信用困境1重复博弈可以归结为图4.20所示的一次性博弈。

		商人2 诚信	欺骗	中
	诚信	7,7	1,6	1,1
商人1	欺骗	6,1	2,2	1,1
	中	1,1	1,1	4,4

图4.20 信用困境(4)

在图4.20中显然存在着3个纯纳什均衡：(诚信，诚信)、(欺骗，欺骗)和(中，中)。根据逆推法，这3个纯纳什均衡都是信用困境1重复博弈的子博弈完美纳什均衡解。图4.20中的纳什均衡(欺骗，欺骗)对应着重复博弈的完美均衡((欺骗，欺骗)，(欺骗，欺骗))，而纳什均衡(中，中)则对应着重复博弈的完美均衡((中，中)，(欺骗，欺骗))。重复博弈中的这两个完美均衡都简单地由两个阶段博弈的纳什均衡组成，但第三个纳什均衡结果却由一个非纳什均衡(第一阶段)和一个纳什均衡(第二阶段)组成。它对应着的重复博弈完美均衡为((诚信，诚信)，(中，中))，与前两个完美均衡本质不同的地方是合作解(诚信，诚信)在第一阶段中出现了。实际上，如果G是一个有着多重纳什均衡的完全信息静态博弈，则重复博弈G(T)可以存在子博弈完美纳什均衡解，其中对每一个$t<T$，t阶段的结果都不是G的纳什均衡。这就从一个方面揭示了为什么合作解能够在重复博弈中出现，因为合作得到奖励，不合作受到惩罚，而且奖励和惩罚机制是一个纳什均衡，即每一方都会加以遵守。但是这种多重纳什均衡的处理仍然存在着令人不满意的地方。因为多重均衡意味着什么事都可能发生，合作解只是一个小概率事件，它的出现需要太多的巧合，而且对于干扰过于敏感，也就是说合作解相当不稳定，这表明子博弈完美对可信性的要求并不严格。例如，在推导子博弈完美解((诚信，诚信)，(中，中))时，我们假定如果第一阶段的结果是(诚信，诚信)，则参与者双方都预期(中，中)将是第二阶段的解，如果

第一阶段出现了任何其他8种结果之一，第二阶段的结果就会是(欺骗，欺骗)。但是，在第二阶段，(中，中)是一个具有优势的纳什均衡，3的收益显然要大于选择(欺骗，欺骗)时只能得1的收益。因而，即使第一阶段有人背叛了诚信，到第二阶段仍选择欺骗就显得愚蠢，因为过去的并不计入成本。一个很自然的想法是，参与双方可能会重新谈判。每一个参与者可能会理性地认为过去的反正已经过去了，在余下的阶段博弈中就会选择双方都偏好的均衡行动(中，中)。但是，如果无论如何最后一个阶段的博弈都为(中，中)的话，那又有什么理由去保证在前面的阶段博弈中去选择(诚信，诚信)。((诚信，诚信)，(中，中))就不再是一个子博弈完美解。

克服上述问题的一个思路就是拒绝谈判，那么为什么拒绝谈判，因为拒绝谈判又比接受谈判有更好的利益。我们考虑图4.21所示的博弈。

商人2

		诚信	欺骗	中	左	右
	诚信	4,4	0,5	0,0	0,0	0,0
	欺骗	5,0	1,1	0,0	0,0	0,0
商人1	中	0,0	0,0	3,3	0,0	0,0
	左	0,0	0,0	0,0	4,1/2	0,0
	右	0,0	0,0	0,0	0,0	1/2,4

图4.21　信用困境(5)

这个阶段博弈在图4.19的基础上又加上了策略左和右，从而阶段博弈有了四个纯策略纳什均衡：(欺骗，欺骗)、(中，中)、(左，左)和(右，右)。与上例相同，和(欺骗，欺骗)相比，参与双方都更倾向于选择(中，中)。但更重要的是，图4.21的博弈中，不存在一个纳什均衡(x,y)，使参与双方和(左，左)或(右，右)或(中，中)相比，都更倾向于选择(x,y)。我们称(中，中)帕累托优于(欺骗，欺骗)，而且(左，左)、(右，右)和(中，中)都处于图4.21所示博弈的纳什均衡收益的帕雷托边界(Pareto frontier)之上。

设想图4.21的阶段博弈重复进行两次，且在第二阶段开始前可以

观测到第一阶段的结果。进一步假设参与者预期的第二阶段结果如下：如果第一阶段的结果为（诚信，诚信），第二阶段将是（中，中）；第一阶段（诚信，x），其中 x 为除诚信之外的任意策略，则（左，左）；第一阶段（x，诚信），其中 x 为除诚信之外的任意策略，则（右，右）；第一阶段（x，y），其中 x，y 为除诚信之外的任何策略，则（中，中）。那么（（诚信，诚信），（中，中））就是重复博弈的子博弈完美解，因为先选诚信，再选中，每个参与者都可得到 4＋3 的收益，但在第一阶段偏离这一选择而选欺骗，却只能得 5＋1/2（选择其他行动收益更低）。更为重要的是，前一例子中遇到的困难在这里并没有出现。基于图 4.21 的两阶段重复博弈，对一个参与者在第一阶段不守信用的惩罚，只能是在第二阶段选择一个对背叛者为劣的纳什均衡。这里的要点是，有三个均衡处于帕累托边界上——其中之一可以奖励参与双方在第一阶段的良好行动，另外两个则可以在惩罚第一阶段不守信用的同时，奖励惩罚者。从而，一旦在第二阶段有必要实施惩罚，惩罚者就不会再考虑选择阶段博弈的其他均衡，于是也就无法说服惩罚者就第二阶段的惩罚行为进行重新谈判。

从上述分析可以看出，图 4.19 和图 4.21 人为设计的痕迹过于明显，毕竟现实生活不是随意设计的游戏。另外一个冲突是，在现实生活中有些博弈只有惟一的纳什均衡，有些甚至没有纳什均衡，但合作解同样出现，而且具有相当的稳定性。这就引出了无限重复博弈。

二、无限重复博弈

合作解要在有限重复博弈中出现要求阶段博弈 G 必须存在多重纳什均衡，但在无限重复博弈中这一条件并不是必需的：即使阶段博弈 G 只存在惟一纳什均衡，无限重复博弈中也可以存在子博弈完美纳什均衡解，其中没有任何一个阶段结果是 G 的纳什均衡。显然这和定理 4.3 相对立，根本的原因就在于博弈可以进行无限期。如果博弈是无限的，那么长远利益就要好于短期利益。

我们仍然考虑信用困境博弈，但它是无限次地进行下去，并且在每一 t 期开始之前，t－1 次阶段博弈的结果都已被观察到。由于明天的一元钱和今天的一元钱的价值是不一样的，最简单的理由是今天的一元

钱如果存入银行那么在明天会变成 $1+r$，所以明天的一元钱只相当于今天的 $1/(1+r)$ 元钱，$1/(1+r)$ 实际上就是经济学中的贴现率。如果未来存在收益序列 $R_1, R_2, R_3, \cdots, R_n$，那么这个未来收益流的现值就为：

$$R_1 + \delta R_2 + \delta^2 R_3 + \cdots = \sum_{t=1}^{n} \delta^{t-1} R_t \qquad (4.12)$$

其中 δ 称为贴现因子。严格地讲，贴现因子并不等于贴现率，但贴现因子与贴现率一定是同方向变动的。例如，我们考虑一个特殊的无限博弈，其结束之前重复进行的次数是随机的，即在博弈的每一阶段完成之后，都要通过抛若干枚（加权的）硬币的方式来决定博弈是否结束，如果硬币朝上那么博弈结束（即概率为 p），如果是其他情况，那么博弈继续（即概率为 $1-p$）。如果下一阶段能得到的收益为 R，那么在当前阶段硬币未抛之前的价值只有 $(1-p)R/(1+r)$；如果下两阶段能得到的收益为 R，在当前阶段硬币未抛之前的价值只有 $(1-p)^2R/(1+r)^2$。令 $\delta=(1-p)/(1+r)$，则贴现因子既包含了货币的时间价值（贴现率），又包含了博弈结束的可能性（$1-p$）。

有了贴现因子的概念，我们就可以比较无限期博弈中的不同收益值，因为无论什么数值（$\neq 0$）在无穷重复相加下它都等于无穷大。例如，在信用困境中，每一阶段得到的收益 4 显然要大于每一阶段获得收益 1，但两者在无限重复博弈中的值之和等于无穷大，即我们根本没有办法对它们进行比较。由于贴现因子 δ 通常位于区间 $(0,1)$ 之间，所以未来收益流的贴现值必然不是无穷数，即无穷级数（4.12）式是收敛的。因而通过贴现因子我们就可以比较无限博弈中不同策略下的收益，知道哪一个策略更好一些。例如，如果每一期的收益为 R，贴现因子为 $1>\delta>0$，那么收益流的贴现值为：

$$\lim_{n\to\infty} R(1+\delta+\delta^2+\cdots+\delta^n)=\frac{R}{1-\delta}$$

下面我们分析无限重复的信用困境博弈，其中每一参与者的贴现因子都为 δ，且每一参与者在重复博弈中得到的收益等于各自在所有

阶段博弈中得到收益的现值。我们将证明尽管阶段博弈中惟一的纳什均衡是不合作，即(欺骗，欺骗)，但在无限重复博弈中子博弈完美纳什均衡却有可能导致这样的结果：每一阶段都是相互合作——(诚信，诚信)。论证中要运用我们分析图 4.20 中的思想，即如果当前参与者选择合作，那么将来(下一阶段)他将得到好报(选择高收益的均衡结果)；如果背叛，将来他将得到报应(选择低收益的均衡结果)。两阶段重复博弈与无限重复博弈本质不同就在于，无限重复博弈下一次选择的高收益均衡，并不是人为加在阶段博弈之上的另外一个均衡，而是代表着在下一阶段及其后的继续合作。

要在无穷多的策略中找出完美均衡是非常困难的，我们不妨先人为主地给出一些我们认为"合理"的策略，然后证明他们是否为完美均衡。其中一个策略就是所谓的触发策略(Trigger strategy)，另一个就是所谓的一报还一报策略。我们首先考虑触发策略，然后再考虑一报还一报策略。

假设参与者 i 在无限重复博弈的开始选择相互合作的策略，并且当且仅当前面每个阶段参与双方都选择相互合作时，在其后的阶段博弈中也选择相互合作。我们可把参与者 i 的这一策略正式表述为：

在第一阶段选择诚信，且在第 t 阶段，如果所有前面 t－1 阶段的结果都是(诚信，诚信)，则选择诚信，否则选择欺骗。

这一策略是触发策略的一种，之所以称为触发策略，是因为如果没有人选择不合作，合作将一直进行下去；一旦有人选择不合作，就会触发其后所有阶段都不再相互合作。如果参与双方都采取这种触发策略，则此无限重复博弈的结果就将是每一阶段选择(诚信，诚信)。我们首先论证如果 δ 距 1 足够近，则采取这种策略，对参与双方而言都是无限重复博弈的纳什均衡，其后再证明这一纳什均衡是子博弈完美的，以使论证更为严格。

要证明采取上述触发策略是无限重复博弈的纳什均衡，等同于证明如果给定参与者 i 的策略为触发策略，那么参与者 j 的最优反应也是触发策略，即触发策略是彼此策略的最优反应。我们将假定参与者 i 已采取触发策略，并证明在 δ 与 1 足够接近的条件下，参与者 j 的最优

反应同样是触发策略。

第一，如果一旦某阶段的结果偏离了（诚信，诚信），参与者 i 将在其后永远选择欺骗；如果某阶段的结果偏离了（诚信，诚信），参与者 j 的最优反应同样是在其后永远选择欺骗。这表明（欺骗，欺骗）是无限重复信用博弈中参与者双方的最优反应。

第二，余下的就是计算参与者 j 在第一阶段的最优反应，以及前面的结果都是（诚信，诚信）时，下一阶段的最优反应。选择欺骗将会使当期得到 5 的收益，但却会触发参与者 i 的永远不合作策略（在对方不合作的情况下，参与者 j 的最优策略只能是不合作），于是未来每一阶段的收益都将成为 1。由于 $\delta+\delta^2+\cdots=\delta/(1-\delta)$，上述一系列收益的现值为：

$$5+\delta\times1+\delta^2\times1+\cdots=5+\frac{\delta}{1-\delta}$$

如果采取合作而不是欺骗，那么本期的收益为 4（虽然小于 5），但在下一阶段还可得到完全相同的选择机会，令 V 表示参与者 j 在（当前和以后每一次面临同样选择时）无限次的选择中总选择最优策略时收益的现值。如果选择诚信是最优的，则：

$$V=4+\delta4+\cdots=4+\delta(4+\delta4+\cdots)=4+\delta V$$

上式清楚地表明，如果当期选择诚信，那么下一期他还可得到完全相同的选择机会。而如果当期选择欺骗，从今以后他惟一的选择只能是欺骗。如果选择欺骗是最优的，则：

$$V=5+\frac{\delta}{1-\delta}$$

于是当且仅当下式成立，选择诚信才是最优反应：

$$\frac{4}{1-\delta}>5+\frac{\delta}{1-\delta} \tag{4.13}$$

即当贴现因子 $\delta>1/4$ 时。① 由此得到结论：当且仅当 $\delta>1/4$ 时，在第一阶段，并且在前面结果都是（诚信，诚信）的下一阶段，参与者 j 的最优反应为诚信。这一结论，再加上前面已证明的，一旦某一阶段的结果偏离了（诚信，诚信），j 的最优反应就是永远选择欺骗。把 j 换成 i，重复上述过程，上述结论同样成立。这就证明了当且仅当贴现因子足够接近 1（$\delta>1/4$）时，参与双方都采取触发策略是无限重复信用博弈的纳什均衡。

第三，现在我们还需要完成最后一步，证明触发策略不仅是无限信用博弈的纳什均衡，而且是子博弈完美纳什均衡。为了证明这一点，我们需要一个辅助性结论。

在有限重复博弈 G(T)和无限重复博弈 G(∞)中，博弈进行到阶段 t 的历史过程(History of play through stages)是指各方参与者从阶段 1 到阶段 t 所有行动的记录。

命题 4.1　在有限重复博弈 G(T)中，由第 t+1 阶段开始的一个子博弈为 G 进行 T−t 的重复博弈，可表示为 G(T−t)。重复博弈中从第 1 阶段起到第 t 阶段为止的每一种可能的历史过程都对应着一个由第 t+1 阶段开始的不同子博弈。在无限重复博弈 G(∞)中，由 t+1 阶段开始的每个子博弈都等同于初始博弈 G(∞)，即 G(∞)=G(∞−t)。与有限重复博弈类似，博弈 G(∞)到阶段 t 为止有多少不同的可能历史过程，就有多少从 t+1 阶段开始的子博弈。

为证明无限重复信用困境中的触发策略纳什均衡是子博弈完美的，我们必须证明触发策略在此无限重复博弈中的每一子博弈中都构成了纳什均衡。根据命题 4.1，无限重复博弈的每一子博弈都等同于原博弈，而触发策略是无限重复信用博弈的纳什均衡，因而它同样是任意一个子博弈的纳什均衡，根据完美均衡的定义可知触发策略是一个完美均衡。这里需要强调的一点是，在无限重复信用困境的触发策略纳什均衡中，当博弈进行到 t 阶段时，存在两个可能的历史过程：(1)所有以前阶段的结果都是（诚信，诚信）的子博弈；(2)至少有一个前面阶段的

① 取 $\delta\geqslant1/4$ 也一样，只不过假设参与者在利益相等下总是选择利他的策略。

结果不是(诚信,诚信)的子博弈。如果参与者在整个博弈中采取触发策略,则:(1)参与者在第一类子博弈(即所有前面的结果都是(诚信,诚信))中的最优策略同样是触发策略,我们已证明它是整个博弈的一个纳什均衡;(2)参与者在第二类子博弈(至少有一个结果不为(诚信,诚信))中的最优策略是永远单纯重复阶段博弈的均衡(欺骗,欺骗),它本身就是阶段博弈 G 的纳什均衡。这就证明了无限重复信用困境中的触发策略纳什均衡是子博弈完美的。

另一个令人惊讶的现象是,虽然信用博弈中只有惟一的纳什均衡,但在无限重复下,它却具有许多的子博弈完美纳什均衡。现在我们证明一报还一报(Tak for tit)策略组合也是无限重复信用博弈的子博弈完美纳什均衡。一报还一报策略的正式表述如下:

在第 0 阶段选择诚信。在 $t>0$ 的任意阶段选择的阶段策略等于对手在 $t-1$ 阶段选择的阶段策略。通俗地说,如果参与者 j 在上一阶段选择诚信,那么参与者 i 在当前阶段选择诚信;如果参与者 j 在上一阶段选择欺骗,那么参与者 i 在当前阶段选择欺骗,要使参与者 i 重新选择诚信,除非参与者 j 在上一阶段自动接受一次惩罚(即在明知参与者 i 选择欺骗的情况下,参与者 j 仍然选择诚信)。

一报还一报策略与触发策略最大的不同就是它对最先欺骗者进行惩罚,然后原谅(如果欺骗者主动选择诚信),而对最先被欺骗者给予补偿(在对手选择诚信下他有一次机会选择欺骗)。在触发策略中,惩罚会永远进行下去,而且对被欺骗者也进行了惩罚,因而一报还一报策略要比触发策略合理得多。

现在我们就证明一报还一报策略(严格来讲,上述策略只是一报还一报策略中的一种而已)是一个完美均衡。如果最优策略是诚信,那么:

$$V=4+\delta V \tag{4.14}$$

如果进行欺骗,则有两种可能情况。一种是永远欺骗下去,那么现值为:

$$V=5+\frac{\delta}{1-\delta} \tag{4.15}$$

前面在证明触发策略均衡的时候，我们已经分析过，如果贴现因子足够接近1，那么永远欺骗下去显然不是最优的策略选择。另一种是开始欺骗，而后后悔，希望重新实现合作。

由于我们考虑的是一报还一报策略，因而博弈进行到 t 阶段时的历史过程具有多种可能，这里我们不妨假设参与者 j 在第 0 阶段选择欺骗，但在第 1 阶段他开始后悔，从而选择诚信，以便在第 2 阶段重新实现合作，当然也可以继续选择欺骗，直到 t 阶段选择诚信，在 t+1 阶段重新实现合作。

那么，如果在第 1 阶段后悔，则现值等于：

$$V=5+\delta\times0+4\delta^2+4\delta^3+\cdots+4\delta^t+\frac{4}{1-\delta} \tag{4.16}$$

即参与者 j 在第 0 阶段选择欺骗，得到 5 的收益，但在第 1 阶段选择诚信（参与者 i 选择欺骗，从面对自己进行惩罚，对参与者 i 进行一次补偿），得到收益 0，到了第 2 阶段重新实现合作。

如果参与者 j 在 t 期才后悔，那么他在 t 期选择诚信，从而在 t+1 阶段重新获得合作机会，则现值为：

$$V=5+\delta+\delta^2+\cdots+0\times\delta^t+\frac{4}{1-\delta} \tag{4.17}$$

第一步，证明(4.16)式大于(4.17)式，贴现因子 δ 必须满足的取值范围，要使(4.16)式大于(4.17)式，就等于证明下列不等式成立：

$$4\delta^2+4\delta^3+\cdots+4\delta^t>\delta+\delta^2+\cdots+\delta^{t-1}$$

$$4\delta^2\,\frac{(1-\delta^{t-1})}{1-\delta}>\delta\,\frac{(1-\delta^{t-1})}{1-\delta}$$

即当贴现因子 $\delta>1/4$ 时，(4.16)式大于(4.17)式，这说明先后悔要比晚后悔为好，因为 $t\geqslant3$，所以上述结论对所有 t 期后悔的策略都成立，即它具有一般性。

第二步，证明后悔要比永远欺骗为好，即(4.16)式大于(4.15)式：

$$5+0+4\,\frac{\delta^2}{1-\delta}>5+\frac{\delta}{1-\delta}$$

当贴现因子为 $\delta>1/4$ 时，参与者 j 选择后悔，以求得重新合作要优于永远欺骗。

第三步，证明永远诚信比欺骗之后再后悔要优，这等于证明(4.14)式大于(4.16)式，即：

$$\frac{4}{1-\delta}>5+0+\frac{4\delta^2}{1-\delta}$$

$$4>5-5\delta+4\delta^2$$

$$0>(1-\delta)(1-4\delta)$$

由于 $1>\delta$，所以当贴现因子 $\delta>1/4$ 时，参与者 j 选择永远诚信是最优的策略。

第四步，我们还有一点必须考虑，即参与者 j 有可能进行多次欺骗，然后多次悔过。比较(4.16)式和(4.17)式可知，早悔过比晚悔过要优，因而我们考虑第 1 期欺骗，第 2 期悔过(选择诚信)，然后第 3 期欺骗，接着再次悔过……上述策略不失一般性，那么现值等于：

$$V=5+0+5\delta^2+0+5\delta^4=\frac{5}{1-\delta^2} \tag{4.18}$$

要使得(4.14)式大于(4.18)式，即永远诚信比朝三暮四的策略要优，贴现因子必须满足的值为：

$$\frac{4}{1-\delta}>\frac{5}{1-\delta^2}$$

即贴现因子 $\delta>1/4$ 时，参与者选择永远诚信要比朝三暮四的策略为优。因而我们可知在给定参与者 i 的策略为一报还一报策略的情况下，一报还一报策略是参与者 j 的最优反应。现在将参与者 j 换成参与者 i 重复以上过程，同样可证明，在给定参与者 j 的策略为一报还一报的情况下，参与者 i 的最优策略也是一报还一报。综上所述，当贴现因子 δ 接近 1(>1/4)，一报还一报策略组合是重复无限信用博弈的纳什均衡。根据命题 4.1 可知，任何重复无限博弈的子博弈都是原博弈，因而一报还一报策略组合也是重复无限信用博弈的子博弈完美均衡。

从上面的例子我们得知当贴现因子充分接近 1，无限重复信用博

弈存在合作解，那么是否所有的无限重复博弈G(∞)都存在合作解？如果是这样，那么我们就可以用无限重复博弈来解释人类社会为什么会出现合作解，尽管这个合作解不是任何阶段博弈的纳什均衡。这就引出了无名氏定理。无名氏定理(因为不知道谁最早证明了它，弗里德曼(Friedman)1971年给出了严格的证明)的核心就在于如果贴现因子充分接近1，那么无限重复博弈就存在合作解的子博弈完美均衡，而这个合作解不是阶段博弈的纳什均衡。

定理4.4　无名氏定理　　在无限重复博弈G(∞)中，令$e=(e_1,\cdots,e_n)$为阶段博弈G的极大极小策略收益，且$x=(x_1,\cdots,x_n)$表示G的任何其他可行收益。如果对每一个参与者i有$x_i>e_i$，则存在一个δ^*，使得对所有的$\delta>\delta^*$，无限重复博弈G(∞)存在一个子博弈完美纳什均衡，其平均收益达到$(x_1,\cdots,x_n)$。在给出无名氏定理的证明之前，对于无名氏定理，有三点需要说明：

(1)极大极小策略收益(Max-min payment)。极大极小策略顾名思义就是从最坏的结果出发，从中选择害处最小的结果的策略。因而极大极小收益是一个博弈的参与者无论如何都能获得的收益(保留收益)。例如在信用博弈中，欺骗就是一个极大极小策略，因为如果选择诚信，那么最坏的结果为0，如果选择欺骗，那么最坏的结果为1，两害相权取其轻，因而欺骗就是极大极小策略，它相对应的收益就为极大极小策略收益。通常，博弈如果存在严格优策略，那么保留收益等于纳什均衡下的收益，如果不存在严格优策略，那么保留收益通常要小于纳什均衡下的收益。见图4.22的斗鸡博弈。其中(让，让)就是极大极小策略(但它不是纳什均衡)，而相应的收益－10就是该博弈的极大极小收益，即该博弈参与者所能得到的保留收益。极大极小策略实际上体现的是一种保守原则，就像斗鸡博弈中的参与者，如果他选择撞，那么对手如果犯错选择撞(按理性原则，对手应当选择让)，他就可能遭受无穷大的损失。而他如果选择让，充其量损失不过为－10，因而选择极大极小策略就是最为安全的策略。

		党徒2 让	党徒2 撞
党徒1	让	−10,−10	−10,+10
	撞	+10,−10	$-\omega,-\omega$

图 4.22　斗鸡博弈

由于保留收益是一个博弈参与者无论如何都能得到的收益，因而即使阶段博弈 G 不存在纳什均衡，无限重复博弈也可能存在合作解的子博弈完美均衡。[①] 另外，阿布勒(Abreu)证明，[②]如果参与者采用最严厉的可信的惩罚措施(纳什均衡不一定是最严厉的惩罚措施，而保留收益对应的策略一定是最严厉的惩罚措施)，那么能够实现比采用纳什均衡收益惩罚措施更高的平均收益。在前面的信用博弈中由于保留收益等于纳什均衡收益，因而(4,4)是无限重复信用博弈所能实现的最高平均收益。

(2)可行收益。我们称一组收益$(x_1,\cdots,x_n)$为阶段博弈 G 的可行收益，如果它是 G 的纯策略收益的凸组合，即$\{r_1,\cdots,r_m\}$为参与者选择纯策略组合下所有可能收益组合的集合(r 为向量)，向量$(a_1,\cdots,a_m)$中的任意一个元素 $a_i\geqslant 0$，并且 $\sum_{i=1}^{m}a_i=1$，那么 $a_1r_1+\cdots+a_mr_m$ 就称之为凸组合。实际上，凸组合等价于参与者以随机的方式来选择纯策略。例如在信用博弈中，有 4 个纯策略收益组合(1,1)、(0,5)、(4,4)和(5,0)，那么 $a_1(1,1)+a_2(0,5)+a_3(4,4)+a_4(5,0)$就是一个凸组合，其中 $a_i\geqslant 0$，$\sum_{i=1}^{4}a_i=1$。所有纯策略收益的凸组合为图 4.23 中的阴影部分。可行收

① 斗鸡博弈本质上是对抗博弈，我们只是用它来说明保留收益。定理 4.4 要求博弈本身不能是完全对抗性的，而是要存在有利于双方的合作(即 $x>e$)，而斗鸡博弈不存在能给双方都带来好处的合作，因而它是完全对抗性的，对它而言，定理 4.4 不成立。像石头、剪刀、布这一博弈，定理 4.4 对它同样不成立。

② Abreu, D. 1988:"On the Theory of Infinitely Repeated Games with Discounting", *Econometrica*, 56:383～396.

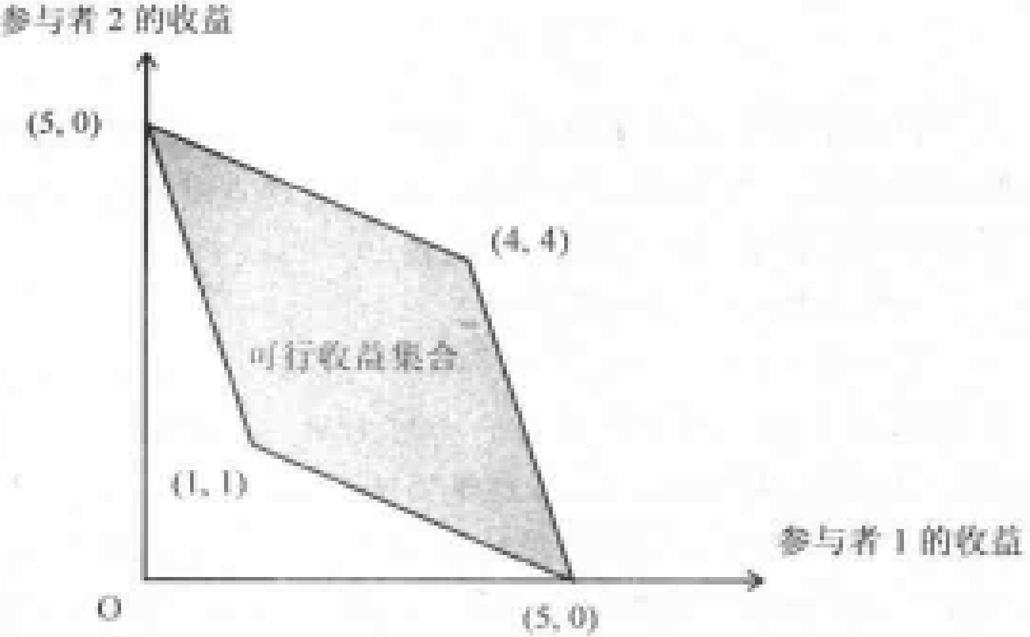


图 4.23　凸组合

益包括 $1<x<4$ 的所有 (x,x)，它们由 $(1,1)$ 与 $(4,4)$ 的加权平均得出 $(a_2=a_4=0)$，以及满足 $y+z=5$ 且 $0<y<5$ 的所有 (y,z)，它们由 $(0,5)$ 和 $(5,0)$ 加权平均得出 $(a_1=a_3=0)$。图 4.23 中阴影区域之内的其他组合是由两个以上的纯策略收益加权平均得出的。另一种理解可行收益的方法就是参与者随机地选择策略，参与者可以使用一个共同的随机发生器，例如根据掷一枚硬币的结果选择(欺骗，诚信)或(诚信，欺骗)，他们可得到的期望收益为 $(2.5,2.5)$。

(3)平均收益。在无限重复博弈 $G(\infty)$ 中，可能每一阶段的收益都不相同，而总的收益值为各阶段博弈收益的贴现值之和，所谓平均收益就是指每一阶段都相同的收益，该收益的贴现值之和等于各阶段博弈收益的贴现值之和。令贴现因子为 δ，$\pi_1,\pi_2,\pi_3,\cdots$ 代表着第一阶段、第二阶段、第三阶段……所有阶段可能的不同收益，它的现值等于 V。如果每一阶段的收益都相等为 π，那么无限重复博弈 $G(\infty)$ 的现值等于 $\pi/(1-\delta)$。如果 $\pi/(1-\delta)=V$，那么 π 就称为无限重复博弈 $G(\infty)$ 的平均收益。

定义 4.8　给定贴现因子 δ，无限的收益序列 $\pi_1,\pi_2,\pi_3,\cdots$ 的平均收益为：

$$\pi = (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\pi_t$$

由于平均收益等于现值 V 的 $(1-\delta)$ 倍，所以使平均收益最大化等同于使现值最大化。使用平均收益的最大优点就在于，我们可以利用它直接和阶段博弈中的收益进行比较。

下面我们证明无名氏定理。

证：已知在无限重复博弈 $G(\infty)$ 中，设 $e=(e_1,\cdots,e_n)$ 为 G 的纳什均衡策略下的收益，并存在一个可行收益 $(x_1,\cdots,x_n)$，并且 $x_i>e_i$，$i\in I$。我们要证明的就是，当贴现因子 $\delta\to 1$ 时，存在一个合作性策略组合（可以为混合策略，即使用随机发生器来决定选择什么行动），使得 $(x_1,\cdots,x_n)$ 成为子博弈完美纳什均衡解的结果。我们把能够实现收益 x_i 的策略称为合作性策略。

我们不妨考虑触发策略，如果所有参与者都采用合作性质的策略，那么在一阶段参与者 i 能得到的收益为 x_i（平均收益）；如果有一个参与者在 t 期背叛了合作性的策略，他在背叛期的收益为 $d_i>x_i$，那么 t 期以后的所有阶段，其他参与者将选择 G 中的纳什均衡。

我们要证明的就是触发策略是无限重复博弈 $G(\infty)$ 的子博弈完美纳什均衡，并且 $(x_1,\cdots,x_n)$ 作为完美解的结果出现。

在给定其他参与者的策略为触发策略的情况下，参与者 i 永远选择合作性策略的现值为：

$$V=\frac{x_i}{1-\delta}$$

如果参与者 i 选择背叛性策略（非合作性的策略），那么根据触发策略，其他参与者将在随后的所有阶段选择 G 中的纳什均衡，则现值为：

$$V=d_i+\frac{e_i\delta}{1-\delta}$$

由于背叛前的收益是一样的，所以我们只要从第一期考虑即可。

要使合作解出现，必须要求：

$$\frac{x_i}{1-\delta}>d_i+\frac{e_i\delta}{1-\delta}$$

整理后得：

$$\delta>\frac{d_i-x_i}{d_i-e_i}$$

也就是说，当参与者 i 的贴现因子充分接近 1，那么参与者的最优选择就是选择合作性策略。当我们对其他参与者 $j\neq i$，重复以上过程，并且取：

$$\delta>\max\{\frac{d_i-x_i}{d_i-e_i}\}_{i\in I}$$

那么触发策略就是无限重复博弈 G(∞)的纳什均衡。

根据命题 4.1 可知，在触发策略下有两种可能的历史过程：第一种是在 t 期以前都是合作性策略，显然触发策略仍是从 t 期开始的子博弈的纳什均衡；第二种是在 t 期之前出现过背叛性策略，显然触发策略也同样是从 t 期开始的子博弈的纳什均衡，因为触发策略中的惩罚性措施是 G 中的纳什均衡。这就证明了触发策略是无限重复博弈 G(∞)中的一个子博弈完美纳什均衡，并且能够实现平均收益($x_1,\cdots,x_n$)。命题得证。

实际上，上述证明中的纳什均衡策略可以改为极大极小策略，相应的收益可以改为保留收益。证明同样成立。这里人们可能会问，极大极小策略是不是一个可置性的威胁，回答是肯定的。因为在明知合作要优于背叛的情况下，参与者 i 还是选择了背叛，作出了损人又损己的事情，因而其他参与者没有理由相信参与者 i 是一个“理性的人”，在疯子什么事都可能作出来的信念下，理性的参与者选择最为保守的策略就是最为安全的策略，这种做法显然是理性的。对参与者 i 不理性的做法的另一种解释就是引入信息不完全，由于信息不完全，参与者 i 就有可能作出了不合理性的选择。例如，在股市上由于信息不对称，当知情人都在抛垃圾股的时候，同样也有大量的人在购买垃圾股。另一种情况是，参与者 i 会利用信息不对称，故意假装非理性已获得更大的利益。

这些信息不对称动态博弈就是我们下一章要讲的主要内容。

接下来为了巩固和丰富所学的知识，我们将举一些经典的例子。

例 4.1　无限重复古诺双头垄断下的共谋。[①]

在第二章中我们曾分析过静态下的古诺双头垄断模型，得出的结论是古诺模型中垄断企业间的共谋(即合作)是不稳定的，共谋下的产量$(a-c)/4$不可能作为均衡结果出现。但在无限重复博弈下，根据无名氏定理可知，当δ接近1时，一定存在一个子博弈完美均衡使得共谋下的产量在无限重复古诺模型中出现。这就解释了为什么现实中会出现反托拉斯法，即使存在反托拉斯法，垄断行业仍然存在普遍的非公开勾结。

我们首先将第二章第三节中的古诺双头垄断模型的主要内容复述如下：市场中的总产量为两个厂商产量之和$Q=q_1+q_2$，市场的总需求曲线为$P(Q)=a-Q$，假定$a>Q$。每一个企业的边际成本为c，固定成本为0，两企业同时选择产量。惟一的纳什均衡为每一企业选择$(a-c)/3$的产量，我们不妨称之为纳什均衡产量，并用q_c表示。由于均衡条件下的总产量$2(a-c)/3$大于卡特尔垄断产量$q_m=(a-c)/2$，如果两企业平均分配卡特尔产量，即$q_i=q_m/2$，那么每一企业能获得较纳什均衡产量更大的利润。

现在考虑上述古诺博弈为阶段博弈的无限重复博弈，两企业的贴现因子均为δ。不妨假设两个垄断企业采取的是触发策略：

在第0阶段生产卡特尔产量的一半，$q_m/2$。第t阶段，如果前面$t-1$个阶段两个企业的产量都为$q_m/2$，则生产$q_m/2$；否则，生产纳什均衡产量q_c。

由于证明过程与前面证明信用博弈的思路完全一样，因而这里不再重复，只把主要的内容简单叙述一遍。

当双方都生产$q_m/2$时，每个企业的利润为$(a-c)^2/8$，我们用π_m来表示。当双方都生产q_c时，每个企业的利润为$(a-c)^2/9$，我们用π_c

① Friedman, J. 1971: "A Non cooperative Equilibrium for Supergames". *Review of Economic Studies*, 38:1～12.

表示。如果企业 i 将在本期生产 $q_m/2$，而企业 j 背叛卡特尔产量，那么本期利润最大化的产量为下式的解：

$$\max_{q_j}(a-q_j-\frac{q_m}{2}-c)q_j$$

它的解为 $q_j=3(a-c)/8$，相应的利润水平为 $9(a-c)^2/64$，我们用 π_d 来表示。那么，要使上述触发策略成为无限重复古诺模型的子博弈完美解，就必须满足：

$$\frac{\pi_m}{1-\delta}>\pi_d+\frac{\delta\pi_c}{1-\delta} \tag{4.19}$$

将 π_m,π_c,π_d 的值分别带入(4.19)式可得 $\delta>9/17$。由此可知，当 $\delta>9/17$时，上述触发策略将是无限重复古诺模型的子博弈完美纳什均衡。

一个有趣的问题是，如果 $\delta\leqslant 9/17$，那么触发策略还能不能维持一个“理想”的产量。由于纳什均衡利润 $\pi_c=(a-c)^2/9$，卡特尔利润 $\pi_m=(a-c)^2/8$，根据无名氏定理，在$[(a-c)^2/9,(a-c)^2/8]$之间存在着许多的可行利润，因而一定存在某些子博弈完美均衡能够实现这些可行利润。见图 4.24 所示。

有两种思路来分析这一问题。第一种思路是，首先计算对任意一个给定的 δ 值，如果双方都采用触发策略，一旦出现背离就永远转到纳什均衡产量，企业可以达到利润最大化的产量。因为当企业的产量小于 $q_m/2$ 时，是一个严格劣策略(想一想，在对手的产量为 0，每个企业利润最大化下的产量是多少？给定对手生产$(a-c)/2$，那么企业的最优产量选择又是多少?)，所以企业的产量选择不可能小于 $q_m/2$。其次对于任意的 δ 值，永远简单重复纳什均衡产量显然是一个子博弈完美纳什均衡。因而，每一个企业所能选择的产量一定位于区间$[q_m/2,q_c]$之内。为了计算这一产量，考虑如下的触发策略：

第 0 阶段生产 q^*。在第 t 阶段，如果在此之前的 t－1 个阶段两企业的产量都是 q^*，则生产 q^*；否则，生产纳什均衡产量 q_c。

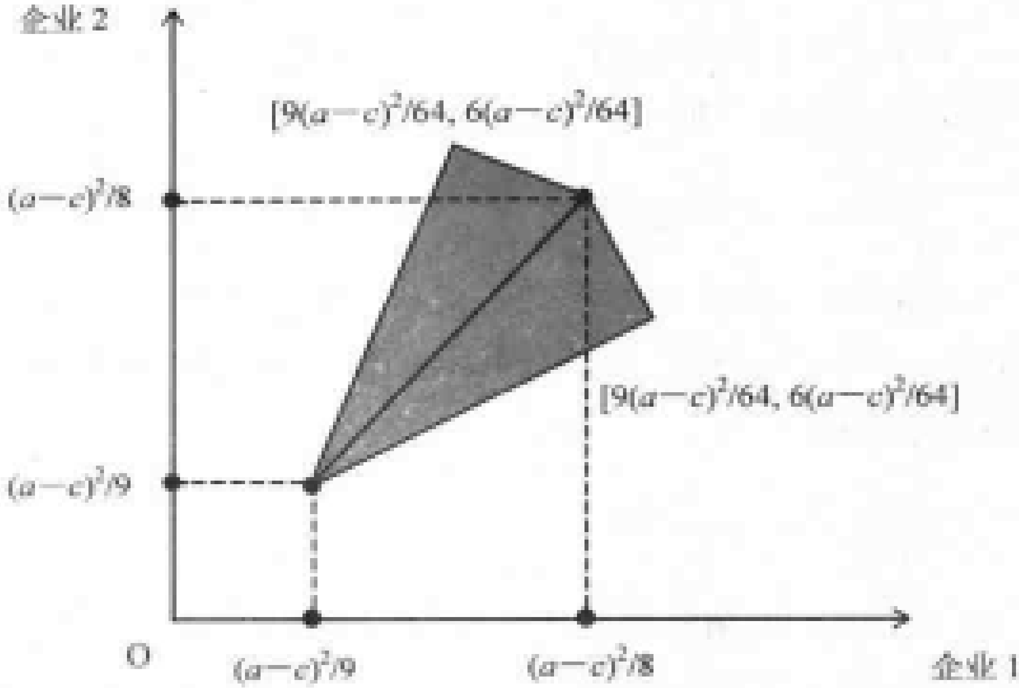


图 4.24 可行利润集合

如果双方都生产 q^*，每个企业的利润为 $(a-2q^*-c)q^*$，我们用 π^* 来表示。如果企业 i 计划在当期生产 q^*，而企业 j 背叛，那么当期利润最大化的产量为下式的解：

$$\max_{q_j}(a-q^*-q_j-c)q_j$$

其解为 $q_j=(a-q^*-c)/2$，相应的利润为 $(a-q^*-c)^2/4$，我们仍用 π_d 表示。当下式成立时，两个企业都采取上面给出的触发策略为完美均衡，即：

$$\frac{\pi^*}{1-\delta}>\pi_d+\frac{\delta\pi_c}{1-\delta}$$

解由此式形成的关于 q^* 的二次方程，可得令上面给出的触发策略成为子博弈完美纳什均衡的 q^*，即：

$$q^*=\frac{9-5\delta}{3(9-\delta)}(a-c)$$

它随 δ 单调递减，且当 $\delta=9/17$ 时，达到 $q_m/2$，当 $\delta=0$ 时达到 q_c。

第二种思路就是企业威胁使用最严厉的可置信惩罚。企业的惩罚

越严厉，那么对 δ 充分接近 1 的要求就越松。这里我们只简单地证明，如果惩罚更为严厉，那么在 $\delta=1/2$ 时，也可以达到卡特尔产量。考虑下面的萝卜加大棒(Carrot and stick)策略：

在第 0 阶段生产卡特尔产量的一半，即 $q_m/2$。第 t 阶段，如果两个企业在第 t－1 阶段都生产 $q_m/2$，则生产 $q_m/2$；如果两个企业在 $t-1$ 阶段的产量都是 x，则生产 $q_m/2$；其他情况下生产 x。

这一策略为参与者提供了两种手段：其一是单阶段的惩罚，这时企业生产 x；其二是潜在无限阶段的合作，这时企业的产量为 $q_m/2$。如果任何一个企业偏离了合作，则惩罚开始。如果任何一个企业背离了惩罚，则会使博弈进入又一轮惩罚。如果两个企业都不背离惩罚，则在下一阶段又回到合作。

如果两企业都生产 x，那么每个企业的利润为 $(a-2x-c)x$，我们用 $\pi(x)$ 表示。令 $V(x)$ 表示当期的利润为 $\pi(x)$，以后每阶段的利润永远是卡特尔利润的一半，企业总收益的现值为：

$$V(x)=\pi(x)+\frac{\delta\pi_m}{1-\delta}$$

如果企业 i 选择产量 x，而企业 j 背叛，那么企业 j 利润最大化时的产量为下式的解：

$$\max_{q_j}(a-q_j-x-c)q_j$$

其解为 $q_j=(a-x-c)/2$，相应的利润为 $(a-x-c)^2/4$，我们用 $\pi_{dp}(x)$ 来表示，其中 dp 的含义是对惩罚的背离。

为了便于大家理解，我们将所有涉及的产量和利润列成表 4.1。

表 4.1　产量和利润表

产　量	利　润
$q_c=(a-c)/3$	$\pi_c=(a-c)^2/9$
$q_m=(a-c)/4$	$\pi_m=(a-c)^2/8$
$q_d=3(a-c)/8$	$\pi_d=9(a-c)^2/64$
x	$\pi(x)=(a-2x-c)x$
$q_{dp}=(a-x-c)/2$	$\pi_{dp}(x)=(a-x-c)^2/4$

如果两个企业都采用上面的萝卜加大棒策略，则无限重复古诺模型中进行到第 t 阶段的历史过程（子博弈）就可归为两类：(1)合作的子博弈，其前面一个阶段的结果是$(q_m/2, q_m/2)$或(x, x)；(2)惩罚的子博弈，其前面一个阶段的结果既非$(q_m/2, q_m/2)$，又不是(x, x)。两企业都采取上面的萝卜加大棒策略，要成为一个子博弈完美纳什均衡，则该策略在每一类子博弈中必须是纳什均衡。具体地说，第一，在合作子博弈中，选择永远合作要优于背叛之后的再次合作，即每一企业获得卡特尔利润的现值要大于本期得到π_d与下期得到惩罚的现值收益$V(x)$之和。

$$\frac{\pi_m}{1-\delta}>\pi_d+\delta V(x) \tag{4.20}$$

将$V(x)$代入(4.20)式可得：

$$(1+\delta)\pi_m>\pi_d+\delta\pi(x)$$

$$\delta[\pi_m-\pi(x)]>\pi_d-\pi_m \tag{4.20.1}$$

它表示，在本期背离所得到的好处必须小于下一期惩罚带来损失的现值。实际上，上述结论非常直观，因为无论是永远合作还是背叛之后的再次合作，从第 2 期开始它们的现值都是一样的，所以是否背叛就取决于本期背叛所得好处是否大于下一期惩罚带来的损失现值。

将表 4.1 中的值相应带入(4.20.1)式，整理后得：

$$8(1+\delta)>9+\frac{64\delta x}{a-c}-\frac{128\delta x^2}{(a-c)^2}$$

令$y=x/(a-c)$，$\delta=1/2$，得：

$$64y^2-32y+3\geq 0$$

解得的结果是：如果$x/(a-c)$不在 1/8 到 3/8 之间，(4.20)式即可满足。

第二，在惩罚子博弈中，企业本期接受惩罚比继续背叛，且下一期继续接受惩罚要好：

$$V(x)>\pi_{dp}(x)+\delta V(x) \tag{4.21}$$

将 $V(x)$代入(4.21)式，得：

$$\delta[\pi_m-\pi(x)]>\pi_{d_p}(x)-\pi(x) \tag{4.21.1}$$

即老老实实接受惩罚要好过继续背叛。我们将表 4.1 中的值代入(4.21.1)，仍令 $\delta=1/2$，解得的结果是：如果 $x/(a-c)$处于 3/10 到1/2 之间，那么(4.21)式必然成立。从而，对贴现因子 $\delta=1/2$ 可达到卡特尔产量的萝卜加大棒策略成为子博弈完美纳什均衡的条件是 $3/8<x/(a-c)<1/2$。

实际上，我们这里考察的古诺模型相对简单，它还可以进一步扩展。例如，引入状态的随机性，这时垄断企业面对的市场需求是随着经济扩展和收缩随机波动；或是引入一定的信息不完全，其中一个重要的模型就是不完美监督模型。下一个例子效率工资就会涉及这一点。

从某种意义上说，萝卜加大棒政策要比触发策略为优，当人们无法完美地观察到企业的背离时，情况就更是如此。如果企业不能观察到彼此的产量选择，而只能观察到市场出清价格，而这一价格在每一阶段又会受到一些随机因素的冲击，例如经济衰退、战争。在这样的条件下，企业就很难对市场出清价格下降是由于其他企业背叛共谋产量，还是由于其他不利因素造成的作出判断。按照触发策略，任何低于触发水平的价格都会引发价格战，并导致永远的惩罚。然而，市场因素一次严重的不利冲击也会使价格降到触发水平，从而导致永远的惩罚，由于惩罚是由于偶然因素引起，因而触发策略中导致惩罚永远进行下去的做法就不再是最优的，显然萝卜加大棒的政策要更为合理。

例 4.2　效率工资。

在新凯恩斯宏观经济学中，效率工资是说明工资刚性和失业并存的重要模型。在效率工资模型中，企业为了激励工人努力工作，一方面支付比劳动力市场均衡工资高的工资，另一方面又威胁如果发现工人偷懒将立即开除。这种做法导致的后果就是工资刚性和人为制造的非自愿失业并存，并且这种状态能够在劳动力市场中长期存在。见图4.25 所示。

显然这种情况在传统经济学中是不可能出现的。然而我们在后面

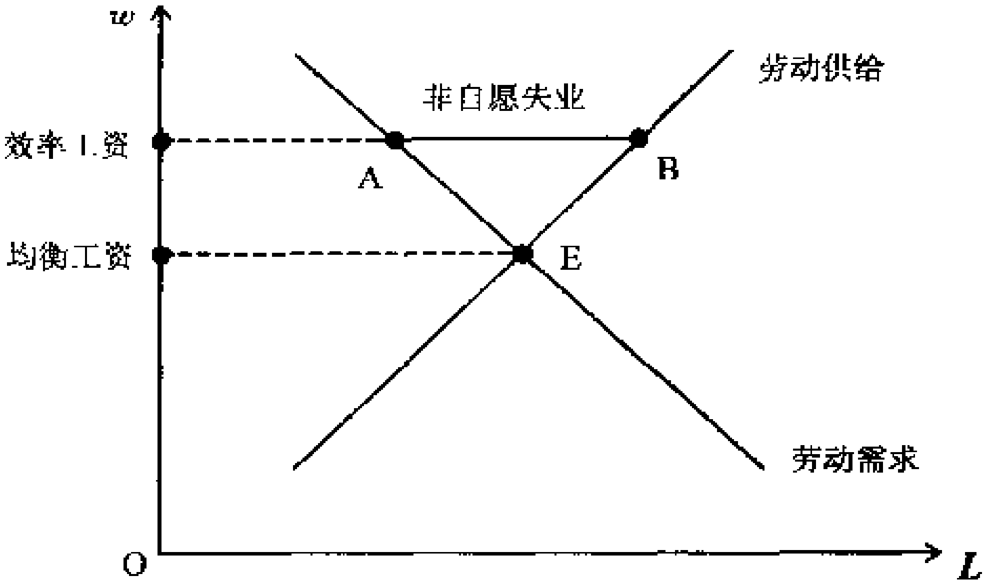


图 4.25 效率工资与失业

将证明，效率工资这种策略（实际上是一种触发策略）在企业和工人组成的无限重复博弈中可以是一个子博弈完美纳什均衡解。效率工资越高，失业工人的人数越多，一个被开除的工人寻找到新的就业机会的可能性就越小，于是开除的威胁就越有效。在博弈均衡条件下，工资水平 w 和失业率 u 恰好可以使工人不敢偷懒，并且企业在工资水平为 w 时的劳动需求恰好使得失业率等于 u。这里我们分析一个最简单的模型，[①]即只包含一个企业和一个工人。

考虑如下的阶段博弈：第一，企业给出一个工资水平 w；第二，工人根据 w 决定接受或拒绝工作。如果工人拒绝 w，则工人成为自我雇用者，工资水平为 w_0，如果工人接受了 w，则工人选择是努力工作（会带来 e 的负效用）还是偷懒（不会带来任何负效用）。工人的努力程度企业无法观测，但企业能观测到工人的产出水平。产出可能高也可能低，为简单起见，不妨假设低产出为 0，高产出为 $y>0$。如果工人努力工作则肯定可以得到高产出，但如果工人偷懒则以 p 的概率得到高产出，$1-p$ 的概率得到低产出。从而，在此模型中，低产出是偷懒无可辩驳的证据。

① Shapiro, C. and Stiglitz, J. 1984:"Equilibrium Unemployment as a Discipline Device", *American Economic Review*, 74:433～444.

假设企业以 w 的工资雇用了工人，那么如果工人努力工作，带来高产出时参与人的收益分别为：企业 $y-w$，工人 $w-e$。如果工人偷懒，则 e 变为 0；如果出现低产出，则 y 变为 0。我们假定 $y-e>w_0>py$，对工人来讲，受雇于企业并且努力工作是有效率的，工人自我雇用要优于受雇于企业并偷懒。

这一阶段博弈的纳什均衡解是企业开出 0 的工资，工人选择自我雇用。工人受雇于企业这样的合作解并不会出现。因为如果企业开出 w 的工资，工人选择偷懒是一个严格优策略（选择努力收益为 $w-e<w-0$ 选择偷懒）。在有限重复博弈中结论同样成立。但如果这个博弈重复进行无穷多次，那么根据无名氏定理，一定存在一个子博弈完美纳什均衡使得合作解，即工人受雇于企业在博弈中出现。而效率工资就是该博弈子博弈完美纳什均衡中的一个。

如果所有前面的工资开价都是 w，所有的开价都被接受了，并且所有前期的产出都是高的，我们就称博弈的过程是"高工资、高产出"。企业的策略为第一阶段开出工资水平 $w=w^*$，并且在其后的每一阶段，如果博弈的过程是高工资、高产出，则继续开出工资水平 w；但其他情况下开出 $w=0$。工人的策略为如果 $w\geq w_0$，则接受企业的工资（否则，选择自我雇用），并且如果博弈的过程（包括本阶段的工资）是高工资、高产出，则努力工作（否则偷懒）。

企业的策略类似于前面分析的触发策略：如果在前面所有阶段的博弈中都相互合作则继续合作，但一旦有一次合作被打破，就永远转向阶段博弈的纳什均衡解。工人的策略也类似于触发策略，但由于工人在阶段博弈中后行动，其策略也更加灵活一些。在一个基于静态阶段博弈的重复博弈中，一方的背离只能在当前阶段结束时才可观测到；但在基于动态阶段博弈的重复博弈中，首先行动方的背离在同一阶段就可被观测到，并可对其作出反应。工人的策略将会是：如果前面所有阶段的博弈都合作则继续合作，如果企业一旦有背离就选择自己本阶段的最优行动，因为他知道将来所有阶段的博弈都将出现阶段博弈的纳什均衡。具体地说，如果 $w\neq w^*$，但 $w\geq w_0$，则工人将接受企业的工资，但选择偷懒。

下面我们将导出上述双方的策略为子博弈完美纳什均衡的条件。同前面一样，论证由两部分组成：(1)导出双方策略成为纳什均衡的条件；(2)证明它们是子博弈完美的。

假设企业在第0阶段开出的工资为w^*。给定企业的策略，工人接受这一工资水平是最优的。如果工人努力工作，则他肯定得到高产出，那么企业将再次开出工资水平w^*，而工人将在下一阶段就努力与否进行相同的决策。因此，如果对工人来讲努力工作是最优的，则工人收益的现值为：

$$V_e = w^* - e + \delta V_e$$

如果工人偷懒，则工人将以p的概率得到高产出，这时下一阶段他还可以就努力与否进行决策；但工人还将以$1-p$的概率得到低产出，这时企业将在以后永远开出工资$w=0$，于是工人亦将永远选择自我雇用。因此，如果对工人来讲偷懒是最优的，则工人的收益现值为：

$$V_s = w^* + \delta\left[pV_s + (1-p)\frac{w_0}{1-\delta}\right]$$

$$V_s = \frac{1}{(1-p\delta)(1-\delta)}\left[(1-\delta)w^* + \delta(1-p)w_0\right]$$

对于工人来讲，要使选择努力为最优必须满足$V_e > V_s$，即

$$w^* \geqslant w_0 + \frac{1-p\delta}{\delta(1-p)}e = w_0 + \left[1 + \frac{1-\delta}{\delta(1-p)}\right]e \tag{4.22}$$

于是，为激励工人努力工作，企业必须向工人支付的w，不仅足以补偿工人自我雇用时的机会收入以及努力工作带来的负效用w_0+e，还包含工资升水$(1-\delta)e/\delta(1-p)$。很自然，如果p接近于1(如果偷懒很难发现)则工资升水必须非常高才可以激励工人努力工作。另一方面，如果$p=0$，则在下式成立时，工人努力工作是最优的：

$$\frac{1}{1-\delta}(w^* - e) \geqslant w^* + \frac{\delta}{1-\delta}w_0 \tag{4.23}$$

上式又可化为下式：

$$w^* \geqslant w_0 + (1 + \frac{1-\delta}{\delta})e$$

它实际上就是(4.22)式中 $p=0$ 时的情景。

即使(4.22)式成立，从而工人的策略为其对企业策略的最优反应，还应该研究企业支付 w 是否值得。给定工人的策略，企业在第 0 阶段的问题可归为就以下几点进行选择：(1)支付 $w=w^*$，并通过威胁工人一旦出现低产出就将其开除来激励工人努力工作，这样每一阶段都可得到 $y-w$ 的收益；(2)支付 $w=0$，促使工人选择自我雇用，自己在每一阶段的收益均为 0。于是，企业策略成为工人策略最优反应的条件为：

$$y - w^* \geqslant 0 \tag{4.24}$$

前面我们假设工人选择受雇于企业并努力工作是有效率的，即 $y-e>w_0$。但要使双方策略成为子博弈完美纳什均衡，还需要进一步的条件，即将(4.22)式和(4.24)式合并得：

$$y - e \geqslant w_0 + \frac{1-\delta}{\delta(1-p)}e$$

这说明要使效率工资能够成为完美均衡，即维持企业和工人的雇佣关系，需要一个充分接近 1 的贴现因子 δ，只有 δ 足够大，这样的合作关系才可能出现。

到此为止，我们已证明如果(4.22)式和(4.23)式成立，则前面给出的策略为纳什均衡。为证明这些策略同时又是子博弈完美的，首先需定义原重复博弈的子博弈。从前面的分析可知，如果阶段博弈为静态博弈，那么重复博弈的子博弈由某个阶段博弈开始，但在阶段博弈为序贯行动的情况下，子博弈不仅可从阶段博弈开始，还可以始于每一阶段之中——在工人观测到企业给出的工资水平之后。给定参与者的策略，我们可以把子博弈归为两类：(1)始于高工资、高产出之后的子博弈；(2)其他历史过程之后的子博弈。我们已证明在高工资、高产出的历史过程中，参与者的上述策略为纳什均衡。剩下的就是要证明后一种历史过程的情况：由于工人将来不再会努力工作，企业促使工人选择自我就业是

最优的；由于企业将在下一阶段及其后永远支付工资 $w=0$，工人在当前阶段也不会努力工作，并只有在 $w>w_0$ 时才会接受给付的工资。

在这一均衡中，自我雇用是永远的：如果工人曾有一次被捉住偷懒，则企业在其后将永远给付工资 $w=0$；如果企业曾偏离 $w=w^*$（最优工资），则工人将不再努力工作，于是企业也不会再雇用这个工人（可以有很多理由对这种永远自我雇用的合理性提出疑问）。一个自然的想法是，在单一企业、单一工人的模型中，和永远选择阶段博弈的纳什均衡解相比，双方参与者都更愿意回到无限重复博弈的高工资、高产出均衡，这就涉及前面介绍过的重新谈判。但问题是，如果能够重新谈判，即参与者知道惩罚将不会被执行，则以这种惩罚相威胁促成的合作便不再是一个均衡。

在存在劳动力市场的情况下，如果企业同时雇用了许多工人，则它更倾向于不进行重新谈判，因为和一个工人的重新谈判会使处于高工资、高产出均衡的其他工人（或将要选择这一均衡的工人）十分失望而改变策略。如果存在许多企业，问题是企业 i 是否会雇用以前企业 i 曾雇用过的工人。合理的结果可能是“否”。因为它担心会使现有高工资、高产出的工人失望，正如一个企业的情况一样。诸如此类的原因可以解释为什么日本大企业间白领男性成年雇员缺乏流动性。

换一种情况，如果被解雇的工人总可以找到比自我雇用更好的工作，则这时那些新工作的工资（减去努力带来的负效用）就起到了自我雇用收入 w_0 的作用（机会成本）。在一个被雇用工人根本不会受到任何损失的极端情况下（例如他有许多工作可做，或是吃“大锅饭”），即使是无限重复博弈也无法提供对偷懒进行有效的惩罚，从而不存在工人将努力工作的子博弈完美纳什均衡。效率工资模型实际上具有一般性，例如在资本市场的债权人—债务人、股市的股东—公司经理人以及公民—政府官员，如果债务人、经理人、政府官员具有许多的机会，他们的命运不是由委托人说了算（债务人复杂的社会关系、经理人的行政任命、政府官员的上面任命），那么代理人努力工作就不会是完美均衡解。

例 4.3　前后一致（Time consistent）的货币政策。

考虑如下序贯行动博弈，其中雇主和工人就名义工资进行谈判，之

后货币当局选择货币供给，货币供给量又决定了通货膨胀率。如果工资合同无法完美地指数化，雇主和工人在决定工资时都将尽力去预测通货膨胀。不过，一个不完全指数化的名义工资一旦设定，真实的通货膨胀率如高于预测的通货膨胀率，将会使工人实际收入下降，导致雇主扩大雇佣人数，扩张生产。这样货币当局就需要在通货膨胀成本和意料之外的通货膨胀（即高于预测水平之上的通货膨胀）使失业率降低，以及提高总产出之间进行权衡。

我们用下面的阶段博弈来分析这一问题的简化形式。①第一，雇主形成一个对通货膨胀的预期值 π^e；第二，货币当局观测到这一预期并选择真实的通货膨胀率 π。雇主的收益为 $-(\pi-\pi^e)^2$，即雇主总是简单地试图正确预测通货膨胀率，在 $\pi=\pi^e$ 时他们达到收益最大化（最大化收益为 0）。货币当局从自身目标出发，希望通货膨胀率为 0，同时产出 y 能达到有效率的水平 y^*。我们可以把货币当局的收益用下式表示：

$$U(\pi,y)=-c\pi^2-(y-y^*)^2$$

其中的参数 $c>0$ 代表了货币当局在两个目标之间的替代关系。假设真实产出可表示为如下的目标产出和意料之外通货膨胀的函数：

$$y=by^*+d(\pi-\pi^e)$$

其中，$b<1$ 反映了产品市场上垄断力量的存在（如果没有意料外的通货膨胀则真实产出小于有效率的产出水平），且 $d>0$ 表示意料之外通货膨胀通过真实工资对产出起作用，正如前面已经提到的。由此我们可以将货币当局的收益重新表示为：

$$W(\pi,\pi^e)=-c\pi^2-[(b-1)y^*+d(\pi-\pi^e)]^2$$

为解出这一阶段博弈的子博弈完美解，首先我们计算对给定的雇主期望的通货膨胀率 π^e，货币当局的最优选择。令 $W(\pi,\pi^e)$ 最大化，即：

$$\max_{\pi} W(\pi,\pi^e)$$

---

① Barro, R. and Gordon, D. 1983:"Rules, Discretion, and Reputation in a Model of Monetary Policy". *Journal of Monetary Economics* 12, pp. 101～121.

解一阶条件，得：

$$\pi^*(\pi^e)=\frac{d}{c+d^2}[(1-b)y^*+d\pi^e] \tag{4.25}$$

由于雇主预测货币当局将选择 π^*，雇主的问题就是选择 π^e，使 $-(\pi^*-\pi^e)^2$ 等于 0，得到：

$$\pi^e=\pi^*(\pi^e)$$

或 $$\pi^e=\frac{d(1-b)}{c}y^*=\pi_s$$

其中 π 的下标 s 表示阶段博弈，π_s 为阶段博弈均衡下导出的理性预期。由此，我们可以等价地说，雇主持有的理性期望值将是货币当局所希望的通货膨胀水平，因为 $\pi^e=\pi^*(\pi^e)$，从而又有 $\pi^e=\pi_s$。当雇主持有的期望值等于 π_s 时，货币当局设定的 π 略高于 π_s 的边际成本刚好抵消意料之外通货膨胀的边际收益。在这一子博弈完美解中，货币当局被预测到要实施通货膨胀，并且事实上也是如此，但如果它承诺不实行通货膨胀政策就可提高其福利水平。事实上，如果雇主持有理性预期(即 $\pi=\pi^e$)，零通货膨胀率将使货币当局的收益达到最大。

现在考虑双方参与者的贴现因子都为 δ 时无限重复博弈的情况。我们将导出双方的下述策略成为子博弈完美纳什均衡的条件，从而使每一阶段 $\pi=\pi^e=0$。在第 1 阶段，雇主们持有预期 $\pi^e=0$，在其后各阶段，如果所有前期的预期 π^e 都为 0，并且所有前期的真实通货膨胀率 π 也都为 0，则持有预期 $\pi^e=0$；否则，雇主持有预期 $\pi^e=\pi_s$。相似地，在当期预期 $\pi^e=0$，所有以前的预期都为 0，且所有以前的真实通货膨胀率 π 也都为 0 时，货币当局选择令 $\pi=0$ 的货币供给；否则，货币当局选定 $\pi=\pi^*(\pi^e)$，它是对雇主期望值 π^e 的最优反应。

假设雇主在第 1 阶段持有的通货膨胀预期为 $\pi^e=0$。给定雇主们的策略(即在雇主们观测到真实的通货膨胀水平之后对其预期的调整方式)，货币当局可以只集中考虑对如下两个方案的选择：(1) $\pi=0$，它将使下一阶段的 $\pi^e=0$，从而使货币当局在下一阶段仍面临同样的选择；(2)从(4.25)式计算得出的阶段最优选择 $\pi=\pi^*(0)$，它将使得此后所

有的 $\pi^e=\pi_s$，这种情况下货币当局将发现此后永远选择 $\pi^e=\pi_s$ 也是最优的。在本期选择 $\pi=0$，可使得每期的收益均为 $W(0,0)$，而在本期选择 $\pi=\pi^*(0)$，可使本期的收益为 $W(\pi^*(0),0)$，但其后每期的收益永远为 $W(\pi_s,\pi_s)$。从而，当下式成立时，货币当局的策略是雇主调整方式的最优反应：

$$\frac{1}{1-\delta}W(0,0)\geqslant W(\pi^*(0),0)+\frac{\delta}{1-\delta}W(\pi_s,\pi_s) \qquad (4.26)$$

通过对上式的简化可得 $\delta\geqslant c/(2c+d)^2$，参数 c 和 d 都起到两方面的效果。例如，在 d 增大时，可使意料之外的通货膨胀对产出影响的效果增大，但同样的原因 d 增大又使得阶段博弈的解 π_s 增大，后者又增大了惩罚对货币当局的效果。类似地，c 的增大使得通货膨胀更为严重，从而使意料之外通货膨胀的吸引力减小，但同时又使 π_s 下降。在两种情况下，后一影响都超过了前一影响，因而贴现因子的临界值要支持这一均衡，$c/(2c+d^2)$ 必须随 d 的增大而递减，随 c 的增大而递增。

至此，我们就证明了如果(4.26)式成立，货币当局的策略就是雇主策略的最优反应。要证明这一组策略为纳什均衡，还须证明后者是前者的最优反应。我们可以看到，根据这一策略，雇主每一期都可得到其最优的可能收益(即为 0)，因此雇主的策略是最优的。证明这一组策略是子博弈完美纳什均衡的程序与前面的分析完全一样，这里不再赘述。

在博弈论中，完全信息动态博弈相对而言仍然是简单的，通过逆推法一般都能在动态博弈中找出子博弈完美纳什均衡，而且这个均衡具有足够的严格性。然而，如果动态博弈存在信息不完全，那么子博弈完美均衡的要求是否还足够强？逆推法是否还依然有效？在一个完美信息博弈中，逆推法可剔除不可置信威胁，保证行动的前后一致。因为每一信息集都是单点信息集，博弈树上每个决策点都分别代表了参与者可能会遇到并须作出反应的情况。从而在扩展式中，通过每个决策点逐级逆向归纳求解的程序，也就可以看成迫使每个参与者考虑实施他可能作出的每一个威胁。但是，在一个非完美信息博弈中，问题就不那么简单了，因为这种博弈包含了至少一个非单一节点的信息集。在扩展式中运用逆推求解，必然是会碰到非单点信息集。如果博弈真的到达了这

个信息集，强迫参与者考虑如何行动，并不等于让其考虑在一个可能发生的情况下他将如何行动，因为如果博弈真的进行到这个信息集，应该行动的参与者并不知道他到底位于哪一个决策点上。在信息不完全的情形下，这种情况复杂化了。

在逆推法中，解决由非单点信息集引起的问题的方法之一，是在扩展式中从后往前逆推求解，直到遇到一个非单点信息集，先暂时不要管它，径直越过该信息集向上，直到到达一个单点信息集。然后，不仅要考虑在这个单点信息集应该行动的参与者在到达此决策点时如何选择，还要同时分析在刚刚被越过的非单点信息集应该行动的参与者将会如何选择。通常而言，这一方法一般都能得到子博弈完美纳什均衡。解决上述问题的另一种方法，是先在扩展式中运用逆推法，直到遇上一个非单点信息集，然后强迫应该在该信息集行动的参与者考虑如果到达了这一信息集将如何选择，而要做到这一点，必须要求该参与者对到达了同一信息集中哪一个决策点有一个概率上的判断。这样的判断当然又要依赖于各参与者在博弈树的更上方可能的行动。所以，运用这一方法，只对博弈树从下往上逆推是无法得到结果的。实际上，上述两种方法都是一种先逆后顺的方法，即在局部逆而在总体顺（因为在不完全信息动态博弈下，只有起始点（即自然行动）始终是单点信息集），这一程序我们称为顺推法。在顺推法下，要求具备更为严格的均衡概念，这个均衡概念就是所谓的完美贝叶斯均衡和序贯均衡，而这是我们将要在下一章讲授的内容。

复习思考题

1. 什么是子博弈完美纳什均衡？它的提出具有什么理论意义？解完全信息动态博弈的基本方法是什么？为什么它和子博弈完美均衡是一致的？

2. 假设两个人谈判分配一个单位为 1 的物品。第一个人首先提出一个分配方案：$(x_1, 1-x_1)$，x_1 为第一个人所得份额。如果第二个人接受，那么博弈就此结束。如果第二个人拒绝，那么轮到第二个人提出分

配方案：$(x_2, 1-x_2)$，也就是说，x 始终为第一个人的分配份额。假设贴现因子为 $1>\delta>0$。请画出这个博弈的扩展式，并将该博弈表达成扩展式。如果博弈进行 n 次结束，子博弈完美均衡是什么？如果进行无穷多次，子博弈完美均衡又是什么？（请参看 Rubinstein，A. 1982："Perfect Equilibrium in a Bargaining Model"，*Econometrica* 50：97～109.）

3. 请你运用（重复无限期的）囚徒困境博弈来说明无名氏定理。

第五章

非完全信息动态博弈

序列均衡是非完全信息动态博弈中的核心概念，克瑞普斯（Kreps）和威尔逊（Wilson）在1982年最早提出，并发展成为博弈论中最一般的均衡概念。序列均衡是对完美贝叶斯均衡的再精炼，而完美贝叶斯均衡则可以理解为是贝叶斯纳什均衡和子博弈完美纳什均衡的综合。此外，序列均衡也是信息经济学的分析基础。

在这一章中，首先我们会给出完美贝叶斯均衡的定义，其次是对完美贝叶斯均衡的进一步精炼——序列均衡，最后是通过一些举例来加深我们对序列均衡的理解。第三部分则主要讨论对序列均衡的再精炼。对于我们分析的非完全信息动态博弈而言，序列均衡的要求有时仍然不够严格。客观而言，直到今天理论界仍没有找到能适用于所有非完全

信息动态博弈的均衡概念，也没有普适的方法来求解非完全信息动态博弈，序列均衡只不过给了我们这种机会，而要求解较复杂的动态博弈的序列均衡，除了技术之外，还需要灵感加运气。

第一节　序列均衡

在这一节中，我们首先从完美贝叶斯均衡人手，然后过渡到序列均衡。完美贝叶斯均衡，实质上是子博弈完美纳什均衡与贝叶斯纳什均衡的综合，它可以分别看作是子博弈完美均衡和贝叶斯均衡的扩展。在非完全信息静态博弈中，我们曾看到参与者如何猜测对手的类型是相当关键的，参与者的策略实际上是具有类型依存性的。类型猜测（又称信念）不同，得到的策略也会不同，这说明参与者对类型的猜测与他的策略具有同样的重要性，而在非完全信息动态博弈中，这一思想得到了进一步的加强。非完全信息动态博弈中的均衡概念不再是策略组合，而是策略＋信念的组合。克瑞普斯和威尔逊的贡献，就在于将信念与策略结合起来，使得人们能够很方便地分析动态博弈中的信息不完全问题。运用子博弈完美纳什均衡的方法，同时加上猜测的设定，从而解决了在不完全信息下如何设定猜测（猜测下的策略必须是最优反应）、如何定义均衡（策略＋猜测的组合）、如何计算均衡（先逆后顺）的问题。

一、完美贝叶斯均衡(Perfect bayesian equilibrium)

所谓完美是指在动态博弈中，整体的纳什均衡也必须是所有子博弈的纳什均衡，但当非单点信息集在某些策略下走到的可能性为 0 时，我们实际上没有办法运用子博弈完美纳什均衡的要求来对这个策略是否在这些信息集中仍是最优反应加以验证，因为子博弈完美纳什均衡没有提供在这种可能性下我们如何行动的指南。通俗地说，当某一个信息集在整体策略走到的概率为 0，且这个信息集为非单点信息集时，我们就没有办法运用贝叶斯法则去估计它们在各节点的概率是多少。在

这种状况下，由于参与者无法计算它们的收益，因此子博弈完美纳什均衡会忽略这个信息集，而不去检查整体策略在这个信息集及其后面的所有信息集中是否为最优反应。一旦如此就会产生很不合理的结果。这里我们要定义一个新的子博弈概念，即泛子博弈(Super subgame)。泛子博弈与子博弈的惟一区别就是泛子博弈的起始点为一个信息集，而不论是否为单点信息集。

定义 5.1 一个扩展式博弈 $\Gamma=\{I,X,A,p(\cdot),k(\cdot),a(\cdot),H,\rho,u\}$ 中的泛子博弈是该博弈的一个子集，其具有下列特征：

(1)它必须从一个信息集(单点或非单点信息集)开始，并包含该信息集内所有决策点的所有后列点，并且只包含以上这些节点。

(2)假如决策点 x 是泛子博弈的一个决策点，那么每一个 $x'\in h(x)$ 同样是该泛子博弈的决策点。也就是说，如果泛子博弈开始于信息集 h，那么对于任意决策点 $y\in h$，$y>x$ 就意味着对所有 $x'\in h(x)$ 而言，$y>x'$，即泛子博弈不能"分割"任何信息集。

现在我们通过图 5.1 所示的博弈来说明子博弈完美纳什均衡存在的问题。

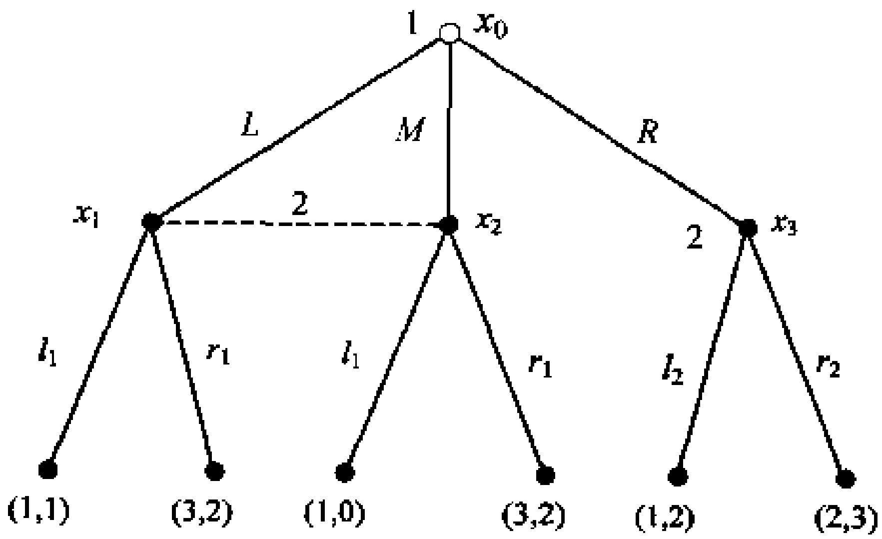


图 5.1 一个例子(1)

图 5.1 所示的博弈，存在一个子博弈完美均衡：参与者 1 选择 R，参与者 2 在信息集 $\{x_3\}$ 选择 r_2，在信息集 $\{x_1,x_2\}$ 选择 l_1。证明的方法如下：给定参与者 2 的策略，若参与者 1 不是选择 R，而是选 M 的话，他

的报酬是 1。若参与者 1 选择 L,他的报酬也为 1,而如果参与者 1 选择 R,那么收益为 2。所以参与者 1 选择 R 是最优反应。反之,在参与者 1 的策略为 R 的情况下,参与者 2 在信息集 $\{x_3\}$ 选择 r_2,在信息集 $\{x_1,x_2\}$ 选择 l_1 的策略是最优反应。因为参与者 2 在信息集 $\{x_3\}$ 的最优策略很明显是 r_2。又由于在参与者 1 的策略为 R 的情况下,$\{x_1,x_2\}$ 这个信息集不可能会走到,因此参与者 2 在信息集 $\{x_1,x_2\}$ 上无论策略是什么都不会影响到他所获得的收益,而且从 $\{x_1,x_2\}$ 出发的这个"博弈"不是子博弈(而是泛子博弈)。这表明策略组合:参与者 1 选择 R,参与者 2 在信息集 $\{x_3\}$ 选择 r_2,在信息集 $\{x_1,x_2\}$ 选择 l_1 既是该博弈的纳什均衡,也是该博弈的子博弈完美均衡。总之,由于原先的策略无法走到信息集 $\{x_1,x_2\}$,所以参与者 2 在此无法计算他在 x_1 或 x_2 的概率各是多少,因此也无法计算他的报酬。因此,子博弈完美纳什均衡无法确定在这个泛子博弈中的策略是否为一个纳什均衡,也因此参与者 2 在信息集 $\{x_1,x_2\}$ 下选择 l_1 符合均衡的要求。

但仔细观察后发现,如果博弈真的进行到了信息集 $\{x_1,x_2\}$,参与者 2 真的会选择 l_1 吗?答案是否定的。因为当参与者 2 走到 $\{x_1,x_2\}$ 时,他选择 r_1 的收益等于 2,而他选择 l_1 则等于 1 或 0,因此不论他在节点 x_1 或 x_2,r_1 一定比 l_1 要好。换言之,在信息集 $\{x_1,x_2\}$ 下,l_1 其实是一个严格劣策略(阶段策略)。出现这种问题的原因是在参与者 1 的策略 R 下不可能走到 $\{x_1,x_2\}$ 这个信息集,所以按照子博弈完美的定义,参与者 2 在 $\{x_1,x_2\}$ 无论选择什么策略都不会影响他的收益。这就是完美均衡中为什么会出现违反理性选择的原因,这表明在动态贝叶斯博弈中只有子博弈完美纳什均衡是不够的,我们需要发展出新的均衡概念。

在完全信息动态博弈中,我们反复强调子博弈完美纳什均衡最大的优点就是可以排除所谓的不可置信威胁。如果我们将这一思想扩展到泛子博弈(而不仅仅是子博弈),那么能不能克服如图 5.1 中出现的问题呢?要将子博弈完美均衡的概念运用到泛子博弈,首先需要解决的问题就是参与者如何作出猜测,这时参与者所谓完整的行动计划不再仅仅包含策略,还要包含对所有信息集概率的猜测,而不管这个信息集能否走到。这就直接导致了完美贝叶斯均衡的产生。

完美贝叶斯均衡的核心就在于它要求即使参与者碰到猜测上的困难，参与者仍应该想办法“制造”出一个对信息集内的各节点的可能猜测。一个“好”的均衡，不但要参与者说明在何处该做什么决定，而且参与者的心中也时时要有一个自己在各个节点的概率猜测。当博弈进行到任意一个节点时，每一个参与者早就胸有成竹——猜测出不同节点的概率各是多少。克瑞普斯和威尔逊的贡献之一就是规定了参与者应当如何“制造”对这些概率的猜测——由博弈原先的设定和参与者所用的策略计算出来，即使有些节点完全不可能走到，参与者也要想办法去设定猜测。因此在每一个参与者的心中，不但要有一个策略存在，还要有一个猜测(Belief)，这个猜测告诉他在不同节点的概率各是多少。这样一个策略加猜测的东西，就是参与者的估算(Assessment)。一个估算实质上就是参与者的万全之策。

定义 5.2 参与者 i 的一个估算是他的策略加上猜测，记为 (σ_i, μ_i)。其中 σ_i 是他的策略，而 μ_i 是一个定义在每一个信息集 $h \in H$ 之上的概率密度函数。也就是说，对每一个 $h \in H$ 和 $x \in h$，$\mu_i(x) \geqslant 0$，且 $\sum_{x \in h} \mu_i(x) = 1$。在任何一个信息集 h 上，μ_i 所代表的意思是，如果博弈现在进行到 h，那么参与者 i 认为他在其中某一个节点 $x \in h$ 的概率为 $\mu_i(x)$。

需要强调的是，参与者的猜测和策略就像一个人的两只手，相互协调实现共同的目标。所谓最优策略是在给定猜测条件下的最优，所谓猜测是在最优策略下的猜测。更精确地说，猜测必须由策略经过贝叶斯法则计算出来，而策略必须是给定猜测下的最优反应。假设 h 是某一信息集，而 (σ, μ) 是一个估算，且 $x \in h$，那么当博弈进行到 h 这个信息集时，在 (σ, μ) 这个估算下，博弈现在正位于 x 点的概率是：

$$\mu(x) = \frac{\pi^{\sigma,\rho}(x)}{\pi^{\sigma,\rho}(h)} \tag{5.1}$$

(5.1)式中的分母，表示在先验猜测 ρ(自然行动的概率)和策略组合 σ 之下，博弈进行到信息集 h 的概率；而分子则代表博弈进行到 x 点的概率。因此当参与者身处 h 时，他认为自己在 x 点的概率 $\mu(x)$，就等

于 $\pi^{\sigma,\rho}(x)/\pi^{\sigma,\rho}(h)$。但是当 h 走到的概率为 0，也就是说 $\pi^{\sigma,\rho}(h)=0$ 时，由于分母为 0，贝叶斯法则无法使用，因此 $\mu(x)$ 无定义。这时对节点 x 的猜测应如何去设定呢？一个最简单的做法是随便设定，但所用的策略必须是这个猜测的最优反应。

定义 5.3 在扩展式博弈 $\Gamma=\{I,X,A,p(\cdot),k(\cdot),a(\cdot),H,\rho,u\}$ 中，若一个估算组合 (σ^*,μ^*) 满足：

(1)在可能情况下，μ^* 一定是从 σ^* 经由贝叶斯法则计算出来的；

(2)若在某一个信息集 h，μ^* 在 h 上的值无法从 σ^* 经贝叶斯法则计算出来，则 μ^* 在 h 上的值可以任意设定；

(3)对每一参与者 $i=k(h)$ 的任一策略 $\sigma_{i=k(h)}$，不等式

$$u_{k(h)}(\sigma^*,\mu^*\mid h)\geqslant u_{k(h)}(\sigma^*_{-k(h)},\sigma_{k(h)},\mu^*\mid h)$$

对每一个信息集 $h\in H$ 成立。

则我们称 (σ^*,μ^*) 为一个完美贝叶斯均衡。

完美贝叶斯均衡所要求的条件，其实相当简单。首先，它要求参与者的猜测必须从策略出发“理性地”加以计算。如果算不出(也就是某个非单点信息集 h 走到的可能性为 0，即 $\pi^{\sigma^*,\rho^*}(h)=0$)，那么猜测可以随意设定。换言之，这时对任何 $x\in h$，$\mu^*(x)$ 的值可以任意猜测，只要 $\sum_{x\in h}\mu^*(x)=1$ 就可以。其次，它要求任一参与者所用的策略，必须是每一个泛子博弈(不只是子博弈)相对于 μ^* 这个猜测的纳什均衡。

我们现在要证明，利用完美贝叶斯均衡的定义，图 5.1 中不合理的子博弈完美纳什均衡 $[R;(l_1,r_2)]$ 并不是一个真正的均衡——完美贝叶斯均衡。

假设 (σ^*,μ^*) 是一个 PBE(完美贝叶斯均衡)，其中 σ^* 是策略组合 $[R;(l_1,r_2)]$。很明显，所有参与者对节点 x_0 和 x_3 的猜测都为 1，即 $\mu_1^*(x_0)=\mu_2^*(x_0)=1$，且 $\mu_1^*(x_3)=\mu_2^*(x_3)=1$。我们惟一要算的是，在信息集 $\{x_1,x_2\}$，$\mu_1^*(x_1)$ 和 $\mu_2^*(x_2)$ 应各是多少。令 $\mu_2(x_1)=t$，$1\geqslant t\geqslant 0$，由于在策略 σ^* 下，$\{x_1,x_2\}$ 走到的可能性为 0，所以 μ_2^* 在 $\{x_1,x_2\}$ 上的值，无法经由贝叶斯法则从 σ^* 计算出来。根据 PBE 的定义，$\mu_2^*(x_1)$ 和

$\mu_2^*(x_2)$的猜测可随意设定，因而任何一个$1\geq t\geq 0$都符合要求。在给定这个猜测下，参与者 2 选择 l_1 的预期收益为：

$$u_2(\sigma^*,\mu^*|\{x_1,x_2\})=t\cdot 1+(1-t)\cdot 0=t$$

选择 r_1 的预期收益为：

$$u_2(\sigma^*,\mu^*|\{x_1,x_2\})=t\cdot 2+(1-t)\cdot 2=2$$

即无论 t 的值是多少，2 都大于 t，因此 l_1 绝对不会是参与者 2 的最优反应。换言之，参与者 2 在$\{x_1,x_2\}$的策略 $\sigma_2^*(x_1,x_2)=l_1$，并不是泛子博弈$\{x_1,x_2\}$的纳什均衡，它和定义 5.3 的条件(3)相矛盾。所以不论对节点 x_1 的猜测 $\mu(x_1)$是多少，σ^* 都不可能是一个完美贝叶斯均衡。

之所以会有这样的结果，原因非常明显。在这个例子中，在策略组合 σ^* 下，信息集$\{x_1,x_2\}$走到的可能性为 0，从而参与者 2 无法根据贝叶斯法则计算他在 x_1 和 x_2 的概率各是多少，而根据子博弈完美纳什均衡的定义，我们并不需要去检查参与者 2 的策略在信息集$\{x_1,x_2\}$中是否为他的最优反应。但 PBE 的要求则进了一步。它要求不管信息集$\{x_1,x_2\}$能否走到，参与者一定要对在 x_1 和 x_2 的可能性做一个猜测。更重要的是，一旦有了这个(不可能走到的信息集的)猜测，他的策略一定要是这个猜测的最优反应。也就是说，他的猜测一定是和策略密切协调的。而在图 5.1 中，由于 l_1 是参与者 2 的严格劣策略，因此不论 $\mu_2(x_1)$或 $\mu_2(x_2)$是多少，他一定会(理性地)选择 r_1，才是对应于 $\mu_2(x_1)$或 $\mu_2(x_2)$的最优反应。也正因如此，图 5.1 中的子博弈完美纳什均衡 σ^* 并不是一个完美贝叶斯均衡。

那么这个博弈的 PBE 是什么？我们可以用逆推法计算。很明显，在 x_3 这个节点，参与者 2 应该选择 r_2。在$\{x_1,x_2\}$这个信息集，假设参与者 2 猜测自己在 x_1 的概率是 $\mu_2(x_1)=t$，而在 x_2 的概率是 $\mu_2(x_2)=1-t$，其中 $1\geq t\geq 0$。若参与者 2 选择 l_1，他的收益为 $t+0(1-t)=t$，若他选择 r_1，那么他的收益为 $2t+2(1-t)=2$。由于 $2>t$，所以参与者 2 一定会选择 r_1。因此不论 t 值为多少，参与者 2 一定会选择 r_1。给定参与者 2

如上的策略，我们再往前看参与者 1 在 x_0 点时，他应该怎么做。如果他选择 L，那么由于他知道参与者 2 接着会选择 r_1，此时他的收益为 3。同理，如果他选择 M，则他的收益也为 3；如果他选择 R，则他的收益为 2。因此参与者 1 随机选择 L 和 M(混和策略)是对参与者 2 策略的最优反应。这意味图 5.1 所示的博弈存在许多的完美贝叶斯均衡：

$\sigma_1(x_0)=(sL,(1-s)M)$，其中 $1\geqslant s\geqslant 0$；$\sigma_2(x_1,x_2)=r_1$，$\sigma_2(x_3)=r_2$。相对应猜测为 $\mu_1(x_0)=\mu_2(x_0)=1$，$\mu_1(x_3)=1$，$\mu_2(x_3)=1$。而 $\mu_2(x_1)=t$，$u_2(x_2)=1-t$，其中 $1\geqslant t\geqslant 0$。

二、序列均衡(Sequential equilibrium)

序列均衡的要求比 PBE 更进一步，它是对完美贝叶斯均衡的再精炼。它要求即使某个信息集在策略下走到的概率为 0，参与者的猜测也不能任意设定，因而序列均衡比起完美贝叶斯均衡要更合理些。在正常的情况下，有些信息集走到的概率为 0。但人非圣贤，孰能无错，一旦参与者无意中出错，走到了原来策略不可能走到的信息集，参与者仍然随意去设定他的猜测吗？显然这样的要求是不合理的。如果一个人端着一杯滚开水，突然手发抖，滚开水泼了一身，发生了烫伤，那怎么办，难道是随意处理吗?这种考虑参与者会犯错误，并在错误发生之后应当如何处理的思想就是序列均衡的核心。因而序列均衡又被称为颤抖手均衡，就好像参与者突然意外地手发抖选错了策略。具体地说，假设某一个参与者有三个可用策略 L、M 和 R。如果他决定选择 R，而不选择其他策略，那么其他参与者仍然会认为他有可能犯错而选择 L 或 M。虽然他选择 L 或 M 的概率很小，但永远大于 0。换言之，选择 R 的这个策略等价于 $(\varepsilon_1 L,\varepsilon_2 M,(1-\varepsilon_1-\varepsilon_2)R)$ 这个策略，其中 ε_1 和 ε_2 无限趋近于 0。这样做的意义是在这种情况下，由于任何一个可用策略被选择的概率都是正的(即使可能性非常小)，因此任何信息集走到的概率一定为正。由于没有一个信息集走到的概率为 0，因此(5.1)式的分母永远不为 0。这表明参与者的猜测永远可根据贝叶斯法则从参与者的策略中估算出来。以下我们就把这个想法严格地定义出来。

定义 5.4　在扩展式博弈 $\Gamma=\{I,X,A,p(\cdot),k(\cdot),a(\cdot),H,\rho,u\}$

中，如果对于所有的信息集 h，有 $\sigma_{k(h)}(a')>0, \forall a' \in a(h)$，$\sigma$ 称为一个完整混合策略(Completely mixed strategy)组合。我们定义 $\sum$ 为扩展式博弈 Γ 所有可能的完整混合策略集合，那么 $\sigma=(\sigma_1,\cdots,\sigma_n)$，且 $\sigma_i \in \sum_i, \forall i \in I$。

完整混合策略的思想在于，博弈树中的任何一个节点被走到的概率永远大于0；在一个完全混合策略组合 σ 下，任一信息集 h 所走到的概率 $\pi^{\rho,\sigma}(h)$ 一定为正。因此给定任何一个估算 (σ,μ)，只要 σ 是完整混合策略组合，那么 $\mu(x)$ 一定可用 σ 经由贝叶斯法则(5.1)式估算出来。亦即：

$$\mu(x)=\frac{\pi^{\sigma,\rho}(x)}{\pi^{\sigma,\rho}(h(x))}$$

定义 5.5 若一个估算 (σ^*,μ^*) 满足下列条件，则称它是一个序列均衡。

(1)对于任一信息集 $h\in H$，任一一个策略 $\sigma_{k(h)} \in \sum_{k(h)}$，不等式

$u_{k(h)}(\sigma^*,\mu^*|h) \geqslant u_{k(h)}(\sigma^*_{-k(h)},\sigma_{k(h)},\mu^*|h)$ 成立。

(2)存在一个序列 $\{(\sigma^t,\mu^t)\}_{t=1}^{\infty}$，使得 $(\sigma^*,\mu^*)=\lim_{t\to\infty}(\sigma^t,\mu^t)$。其中每一个 σ^t 都是完整混合策略组合，且 μ^t 是由 σ^t 经贝叶斯法则计算得来。也就是说，每一个 μ^t 和 σ^t 之间，都必须满足(5.1)式。[①]

从这个定义可以看出，所谓序列均衡其实是一个完整混合策略序列的极限，即它假设所有参与者只使用完整混合策略，但某些行为策略被使用的概率趋近于0，并且完整混合策略的极限必须满足理性的要求(即定义 5.5 中的(1)式)。需要注意的是，我们只要求 (σ^*,μ^*)（无穷序列 $\{(\sigma^t,\mu^t)\}_{t=1}^{\infty}$ 的极限）满足定义 5.5 中的(1)式，而没有要求序列 (σ^t,μ^t) 也必须满足定义 5.5 中的(1)式。这意味着对 $\{(\sigma^t,\mu^t)\}_{t=1}^{\infty}$ 这个序列的选择弹性非常大，这构成了序列均衡的一个缺点，有了后来学者对

① 序列均衡只要求 σ_i^* 是相对于 σ_{-i}^* 的最优反应，而不要求 σ_i^t 是 σ_{-i}^t 的最优反应。换言之，(σ^t,μ^t) 不需要满足定义 5.3 中的(3)式。

序列均衡的再精炼。另外一个值得注意的是，由于(σ^*,μ^*)是无穷序列$\{(\sigma^t,\mu^t)\}_{t=1}^{\infty}$的极限，因此$\sigma^*$可以不是完全混合策略组合。

从上面的对完美贝氏均衡和序列均衡的定义可以看出，它们最大的差异就是对参与者猜测(Belief)的设定。不过幸运的是，在绝大多数情况下，完美贝氏均衡和序列均衡完全一样，两者不同倒是一个特例。现举例如下：

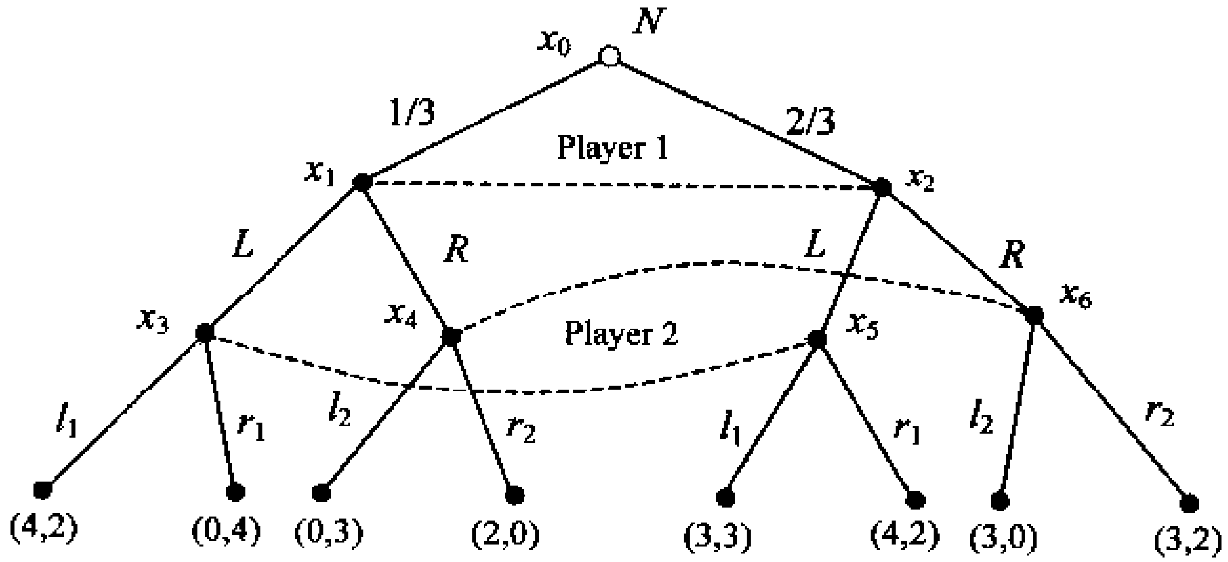


图 5.2　非完全信息博弈

在图 5.2 的博弈中，自然(Nature)选择x_1的概率为 1/3，即$\rho(x_1)=1/3$，选择x_2的概率为 2/3，即$\rho(x_2)=2/3$；参与者 1 可以选择L或R，但是他不确定自己会在x_1或者x_2；参与者 2 亦不知道参与者 1 会从x_1和x_2中哪一点出发。假设有一个均衡策略σ^*是让$\sigma_1^*(L)=1$(表示与参与者 1 以概率 1 选择L)，这意味着$\pi^{\sigma,\rho}(\{x_4,x_6\})=0$，那么在完美贝叶斯均衡的定义下，由于$\{x_4,x_6\}$这个信息集走到的概率为 0，所以参与者 2 在信息集$\{x_4,x_6\}$的猜测，$\mu_2(x_4)$和$\mu_2(x_6)$可以是 0 到 1 之间的任意数，只要参与者 2 在信息集$\{x_4,x_6\}$的策略是相对于$\mu_2(x_4)$和$\mu_2(x_6)$的最优反应即可。但在序列均衡的定义下，就没有这么大的自由度。假设$\sigma_1^t(L)=1-\epsilon^t$，$\sigma_1^t(R)=\epsilon^t$，$1>\epsilon^t>0$，对于$\forall\ t$，且$\lim_{t\to\infty}\epsilon^t=0$。如此一来，$\sigma_1^*$就是$\sigma_1^t$的极限。在这个设定下，有：

$$\mu_2^t(x_4)=\frac{1/3\varepsilon^t}{1/3\varepsilon^t+2/3\varepsilon^t}=1/3,\forall\ t$$

所以 $\mu_2(x_4)=\lim_{t\to\infty}\mu_2^t(x_4)=1/3$。这表明，在序列均衡的设定下，如果 $\sigma_1^*(L)=1$，那么 $\mu_2^*(x_4)$ 非是 1/3 不可。

通过这个例子，我们看出序列均衡与完美贝氏均衡的不同在于，当某一信息集在策略组合 σ^* 下走到的概率为 0 时，完美贝叶斯均衡可以随意设定猜测，而序列均衡则要求尊重原先博弈的信息结构(Information structure)，在参与者猜测的设定上，做了如果出错，仍然理性这样的限制。但也因如此，造成了计算序列均衡较为繁琐的结果。所幸这两种均衡概念，在一般的情形下是完全一样的。我们可以用图 5.3 来加以说明。

假设有一个均衡策略，让 $\sigma_1^*(x_1)=R$，因此 $\pi^{\sigma,\rho}(\{x_2,x_3\})=0$。根据完美贝叶斯均衡的定义，$\mu_2(x_2)$ 可以为任一个 0 到 1 之间的数。而现在我们则根据序列均衡的定义，看一看对 μ_2 有些什么限制。令 $\sigma_1^t(R)=1-\varepsilon_1^t-\varepsilon_2^t$，$\sigma_1^t(L)=\varepsilon_1^t$，$\sigma_1^t(M)=\varepsilon_2^t$，且 $\lim_{t\to\infty}\varepsilon_1^t=\lim_{t\to\infty}\varepsilon_2^t=0$，则 $\lim_{t\to\infty}\sigma_1^t=\sigma_1^*$，并且

$$\mu_2^t(x_2)=\frac{\varepsilon_1^t}{\varepsilon_2^t+\varepsilon_1^t};\mu_2^t(x_3)=\frac{\varepsilon_2^t}{\varepsilon_2^t+\varepsilon_1^t}。$$

对任意一个 $1\geqslant s\geqslant 0$，我们令 $\varepsilon_1^t=\frac{s}{1-s}\varepsilon_2^t$，那么 $\mu_2^t(x_2)=1-s$，$\mu_2^t(x_3)=s$。这说明，在这个例子中序列均衡与完美贝叶斯均衡是完全一样的。其原因在于，虽然序列均衡要求用一个无穷序列去逼近均衡策略，但是无穷序列本身的选择却非常随意，这种随意与完美贝叶斯均衡中的随意设定猜测在绝大多数情况下是完全一样的。因而我们说序列均衡与完美贝叶斯均衡相同是常例，不相同倒是特例。这一点可以帮助我们最大程度地简化对序列均衡的计算。下面我们仍用图 5.3 来说明序列均衡的算法。

仍然使用逆推法。我们不妨假设参与者 2 在信息集 $\{x_2,x_3\}$ 的猜测为 $\mu_2(x_2)=t$，$1\geqslant t\geqslant 0$，那么参与者 2 选择 l 的预期收益为 $2\mu_2(x_2)+$

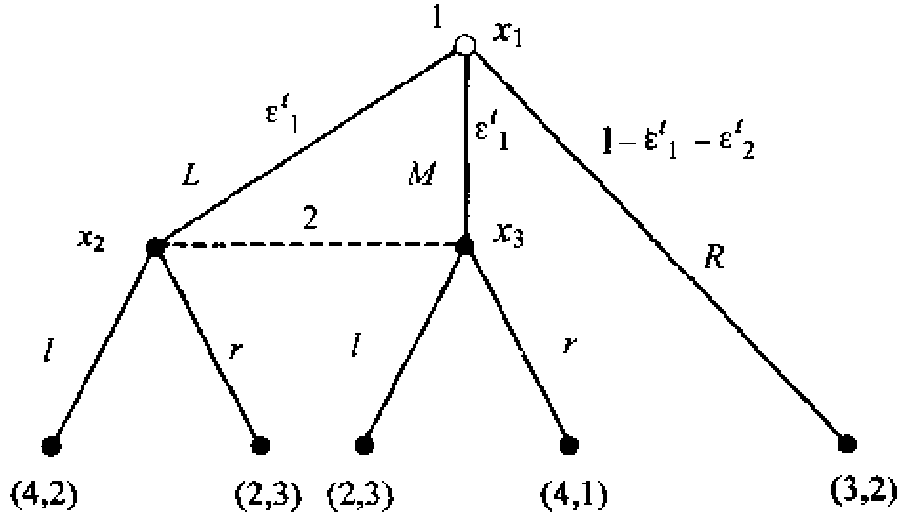


图 5.3　一个例子(2)

$3\mu_2(x_3)=3-t$，选择 r 的预期收益为 $3\mu_2(x_2)+1\mu_2(x_3)=1+2t$。所以当 $t<2/3$ 时，参与者 2 的最优反应是 l；当 $t>2/3$ 时，最优反应是 r；当 $t=2/3$ 时，他的最优反应为混合策略 $(sl,(1-s)r)$，$1\geqslant s\geqslant 0$。

给定参与者 2 上述策略，我们回溯求参与者 1 的最优反应。

若 $t<2/3$，那么由于参与者 2 会选择 l，因此参与者 1 的最优反应是选择 L。但若他选择 L，那么 t 就等于 1，而不可能小于 2/3(这表明猜测和策略不一致)。若 $t>2/3$，那么由于参与者 2 会选择 r，因此参与者 1 的最优反应是选择 M。但若如此，t 就应该等于 0，而不会大于 2/3。最后剩下 $t=2/3$，当 $t=2/3$ 时，参与者 2 的策略为混合策略 $(sl,(1-s)r)$。而让 $t=2/3$ 的可能性有两种：第一，参与者 1 的策略也是混合策略，为 $(2\varepsilon,\varepsilon,1-3\varepsilon)$，① $1/3\geqslant\varepsilon\geqslant 0$；第二，参与者 1 选择 R，而参与者 2 在均衡外信息集 $\{x_2,x_3\}$ 的猜测为 $t=2/3$。前者必须使参与者 1 选择 L 或 M 的预期收益等于选择 R 的收益，这时 L 和 M 的收益必须一样(如果不一样会如何)，这时只有 $s=1/2$ 才能做到。后者必须让参与者 1 选择 L 和 M 的预期收益小于选择 R 的收益，它的条件也是 $s=1/2$。综合二者，可知这个博弈的所有序列均衡为：

$$\sigma_1(x_1)=(2\varepsilon,\varepsilon,1-3\varepsilon),0\leqslant\varepsilon\leqslant 1/3$$

① $t=2/3=2\varepsilon/(2\varepsilon+\varepsilon)$。

$\sigma_2(\{x_2,x_3\})=(1/2l,1/2r)$，且 $\mu_2(x_2)=2/3$，$\mu_2(x_3)=1/3$。

下面我们再考虑对图 5.3 的一个变形。见图 5.4 所示博弈。

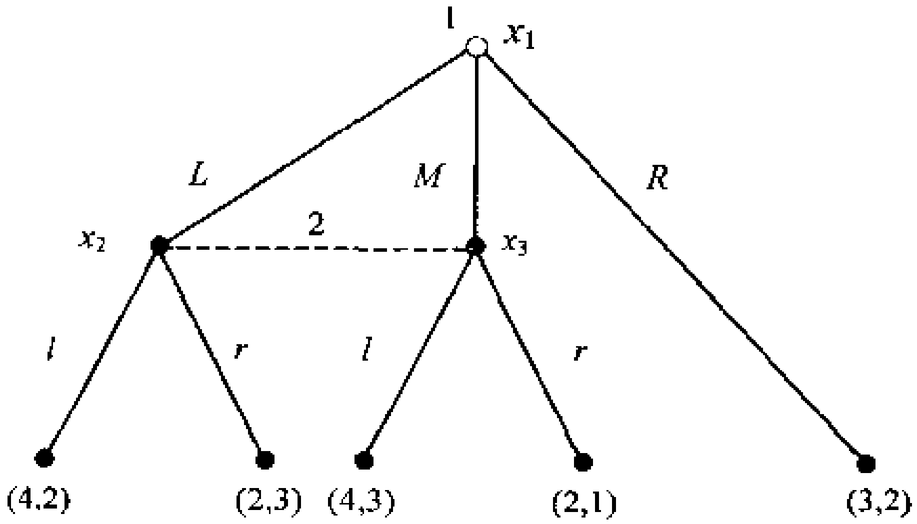


图 5.4 一个例子(3)

我们运用先逆后顺的方法。假设参与者 2 在 $\{x_2,x_3\}$ 的猜测为 $\mu_2(x_2)=t$，$\mu_2(x_3)=1-t$，则他选择 l 时获得的预期收益是 $2t+3(1-t)=3-t$，而选择 r 的预期收益是 $3t+1-t=1+2t$。因此，当 $t<2/3$ 时，参与者 2 会选择 l；当 $t>2/3$ 时，参与者 2 会选择 r；当 $t=2/3$ 时，由于 l 和 r 两者预期收益相等，参与者 2 会选择混合策略 $vl+(1-v)r$，$1\geqslant v\geqslant 0$。我们将讨论分成两部分，其重点在于参与者 1 是否会用 100%的概率选择 R。因为只有参与者 1 用 100%的概率选择 R 时，参与者 2 才会需要在 $\{x_2,x_3\}$ 这个信息集自己做猜测，在其他情况下只要用贝叶斯法则去计算就可以了。

(1)参与者 1 假如未用 100%的概率选择 R，则假设参与者 1 在 x_1 点上的策略是 $\sigma_1(x_1)=(s_1L,s_2M,(1-s_1-s_2)R)$，其中 $1>s_1,s_2>0$，$1>s_1+s_2>0$。

①若 $s_1<2s_2$，则 $\mu_2(x_2)<2/3$，参与者 2 选择 l。这时参与者 1 的收益是 $4s_1+4s_2+3(1-s_1-s_2)=3+s_1+s_2$。因此让 $s_1+s_2=1$ 会使参与者 1 的收益最大，等于 4。因此一个均衡是：$\sigma_1(x_1)=(sL,(1-s)M)$，$s<2/3$，$\mu_2(\{x_2\})<2/3$，$\delta_2\{x_2,x_3\}=l$。

②若 $s_1>2s_2$，则 $\mu_2(x_2)>2/3$，参与者 2 会选择 r。这时参与者 1 的收益为 $2s_1+2s_2+3(1-s_1-s_2)=1-(s_1+s_2)$。当 $s_1+s_2=0$ 时，参与者 1 的收益最大。但这意味着参与者用 100% 的概率选择 R，和我们的假设不符，因此 $s_1>2s_2$ 时没有均衡。

③若 $s_1=2s_2$，则参与者 2 的策略为 $(vl,(1-v)r)$，$1\geqslant v\geqslant 0$，而参与者 1 的预期收益是 $s_1(4v+2(1-v))+s_2(4v+2(1-v)+3(1-s_1-s_2))=3+3s_2(2v-1)>3$。若 $v>1/2$，则 s_2 越大，参与者 1 的收益越高。但由于 $s_1=2s_2$，所以 $s_1=2/3$，$s_2=1/3$ 时，参与者 1 的收益最大。因此有一个均衡：$\sigma_1(x_1)=(2/3L,1/3M)$，$\mu_2(x_2)=2/3$，$\mu_2(x_3)=1/3$，$\sigma_2(\{x_2,x_3\})=(vl,(1-v)r)$，$1>v>1/2$。若 $v<1/2$，则 $s_1=s_2=0$ 时，参与者 1 的收益最大，但这又和参与者 1 不用 100% 的概率选择 R 的假设不符，因此，$v<1/2$ 没有均衡存在。若 $v=1/2$，无论 s_2 等于多少，参与者 1 的收益都是 3。这个情形下的均衡是：$\sigma_1(x_1)=(2sL,sM,(1-3s)R)$，$0\leqslant s\leqslant 1/3$，$\mu_2(x_2)=2/3$，$\mu_2(x_3)=1/3$，$\sigma_1(\{x_2,x_3\})=(1/2l,1/2r)$。

(2)假如参与者 1 用 100% 的概率选择 R，那么参与者 2 便需要预测自己在信息集 $\{x_2,x_3\}$ 下会做什么决定。当 $t<2/3$ 时，参与者 2 应选择 l。但在 $t<2/3$ 时，参与者 1 悖离到 L 或 M 的收益 4 会大于选择 R 的收益 3，因此选择 R 不是参与者 1 的最优反应。当 $t>2/3$ 时，参与者 2 会选择 r，因此参与者 1 选择 R 的收益 3 大于选择 L 或 M 的收益 2。所以 R 是参与者 1 的最优反应。这表明：$\sigma_1(x_1)=R$，$\sigma_2(\{x_2,x_3\})=l$，$\mu_2(x_2)=t$，$t>2/3$，且 $\mu_2(x_3)=1-\mu_2(x_2)$ 是一组序列均衡。当 $t=2/3$ 时，参与者 2 会采用混合策略 $(sl,(1-s)r)$，因为参与者 1 选择 L 的预期收益等于选择 M 的预期收益，为 $4s+2(1-s)=2+2s$。由于选择 R 必须是参与者 1 的最优反应，因此 $2+2s<3$，即 $s<1/2$。因此 $\sigma_1(x_1)=R$，$\sigma_2(\{x_2,x_3\})=(sl,(1-s)r)$，$1/2\geqslant s\geqslant 0$，$\mu(x_2)=2/3$，$\mu(x_3)=1/3$，也是一组序列均衡。

为了便于理解，我们将前面所有可能出现的情况列成表 5.1。通过表 5.1 可以很容易地找出图 5.4 所示博弈的全部序列均衡。

表 5.1　序列均衡计算表

<table>
<tr><td colspan="5">当 $s_1>0, s_2>0, 1-s_1-s_2>0$ 时</td></tr>
<tr><td rowspan="2">$s_1<2s_2$</td><td rowspan="2">$s_1>2s_2$</td><td colspan="3">$s_1=2s_2$</td></tr>
<tr><td>$v>1/2$</td><td>$v=1/2$</td><td>$v<1/2$</td></tr>
<tr><td>$\sigma_1(x_1)$
$=(sL,(1-s)M)$
$s<2/3$
$\mu_2(\{x_2\})<2/3$
$\sigma_2(\{x_2,x_3\})=l$</td><td>没有均衡</td><td>$\sigma_1(x_1)$
$=(2/3L, 1/3M)$
$\mu_2(x_2)=2/3$
$\mu_2(x_3)=1/3$
$\sigma_2(\{x_2,x_3\})$
$=(vl,(1-v)r)$
$1>v>1/2$</td><td>$\sigma_1(x_1)=(2sL,sM,$
$(1-3s)R)$
$0\leqslant s\leqslant 1/3$
$\mu_2(x_2)=2/3$
$\mu_2(x_3)=1/3$
$\sigma_2(\{x_2, x_3\})$
$=(1/2l, 1/2r)$</td><td>没有均衡</td></tr>
<tr><td colspan="5">当 $s_1=s_2=0$，参与者 1 以 100%的概率选择 R 时</td></tr>
<tr><td>$t<2/3$</td><td colspan="2">$t=2/3$</td><td colspan="2">$t>2/3$</td></tr>
<tr><td>矛盾</td><td colspan="2">$\sigma_1(x_1)=R$
$\sigma_2(\{x_2, x_3\})=(sl,(1-s)r)$
$1/2\geqslant s\geqslant 0$
$\mu(x_2)=2/3, \mu(x_3)=1/3$</td><td colspan="2">$\sigma_1(x_1)=R$
$\sigma_2(\{x_2, x_3\})=r$
$\mu_2(x_2)=t, t>2/3$
$\mu_2(x_3)=1-\mu_2(x_2)$</td></tr>
</table>

根据表 5.1，这个博弈所有的序列均衡为：

(1)$\sigma_1(x_1)=(sL,(1-s)M), 2/3\geqslant s, 2/3\geqslant\mu_2(x_2), \mu_2(x_3)\geqslant 1/3,$
$\sigma_2(\{x_2,x_3\})=(vl,(1-v)r), 1\geqslant v\geqslant 1/2$；

(2)$\sigma_1(x_1)=R, \sigma_2(\{x_2,x_3\})=l, \mu_2(x_2)=t, t>2/3$；

(3)$\sigma_1(x_1)=R, \sigma_2(\{x_2,x_3\})=(sl,(1-s)r), 1/2\geqslant s\geqslant 0,$
$\mu_2(x_2)=2/3$。

通过上面的计算可以看出，如果非完全信息动态博弈稍微复杂，计算序列均衡就非常困难，因而计算序列均衡除了技术之外，还需要一点灵感加运气。

第二节　阻吓性定价和诉讼

这一节我们主要给出两个经典模型——阻吓性定价（Predatory

pricing)和诉讼(Litigation)。一方面表明博弈论的运用前景十分广阔，另一方面也表明如何计算序列均衡。

需要指出的是，序列均衡作为博弈论的核心均衡概念，在运用中体现出了巨大的理论威力，然而，它仍然存在着令人不满意的地方。首先，序列均衡的计算十分繁琐和困难，尤其当博弈存在信息不完全和具有动态特征时更是如此。更为要命的是，如何求解扩展式博弈的序列均衡至今仍没有一般通行的方法，常常需要先把均衡猜出来，然后再检验这个猜测是否正确。这就是为什么博弈论发展到今天仍只能研究和处理一些非常简单的模型。

其次，通常求解出来的博弈均衡数非常多，而这些均衡的性质常常不同，甚至相反，这就给我们对博弈的均衡作出解释带来了困难。一个自然的想法是，我们可以“凭感觉”来选择我们认为“合理”或“合乎直觉”的均衡解来讨论，但这是完全没有客观标准的。一个补救的方法，就是加强均衡条件的限制，要求一个均衡不但要满足序列均衡的限制，而且要满足一些更强的条件，从而剔除大量的均衡，达到减少均衡数目的目的。这就是现在博弈论研究中的均衡再精炼(Equilibrium refinemement)。我们在第三节会专门讨论它。

一、阻吓性定价

在初级微观经济学中，垄断厂商的定价是以获得垄断利润最大为原则，其中一个重要条件就是垄断厂商能够阻止新厂商进入该行业。如果不能做到，垄断利润就会吸引新厂商加入该行业，使得垄断利润被打破。形成进入障碍的可能性有专利、政府特许，但也有许多的垄断形成是源自垄断厂商的自然力量。这种自然力量通常有两种做法：一是实行所谓的限制性定价(Limit pricing)。垄断厂商通过制定一个较低的价格，让潜在的进入者认为该行业无利可图，从面打消进入该行业的企图。限制性定价通常在两种情况下产生：一是进入该行业需要巨额的固定投资，例如汽车、航空、石油等行业；或是潜在进入者对该行业内某个影响利润的因素(例如生产的平均成本或市场需求)具有不确定性。二是潜在进入者虽然知道垄断行业存在超额利润，但却迫于在位厂商的

威胁，不敢贸然进入。因为在位厂商威胁一旦有新进入者，它将把价格压至最低，进行价格战，让进入者（和自己）遭受巨大损失。这种策略就是所谓的阻吓性定价。

考虑如图 5.5 一个非常简单的进入模型。

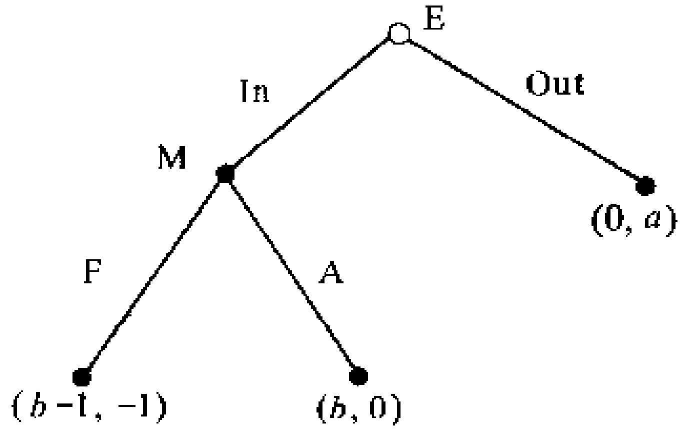


图 5.5　进入模型(1)

新厂商（Enter）决定要不要进入垄断市场。若它选择不进入（Out），则市场内原来的垄断者（Monopoly）得到垄断利润 $a>1$，而 E 的垄断利润为 0。若 E 选择进入垄断行业（In），则垄断者将面临是否进行价格战的选择。如果 M 选择和平共处（Accept），则由于垄断利润被瓜分，所以假设 M 的垄断利润降为 0，而 E 能获得 b 的垄断利润，其中 $1>b>0$。如果 M 选择进行价格战（Fight），则利润为 -1，而 E 得到的利润为 $(b-1)$。可知该博弈惟一的纳什均衡是新厂商选择进入（In），而垄断厂商选择和平共处（A）。如果进入博弈重复进行 T 期（不是无限期），那么根据我们在前面分析重复博弈时得出的结论（定理 4.3），在垄断厂商的威胁是不可置信的情况下，即使博弈进行多期，惟一的子博弈完美纳什均衡仍然是新厂商每一期都进入，而垄断厂商选择和新进入者和平共处，即阶段博弈的纳什均衡在动态博弈中重复 T 次。然而现实与上述理论分析却不一致，我们通常观察到的现象是垄断厂商往往能够有效阻吓新厂商的进入。一个解决的办法是引入无限期，由于博弈重复进行无限期，因而垄断厂商为了长期利润就会不惜代价发动价格战，从而有效阻吓新厂商的进入（无名氏定理）。但现实中仍有大量事实表明，即使博弈重复有限期，垄断厂商在新厂商进入垄断行业后仍会

发动价格战，从而有效阻吓进入者，这种理论与现实的矛盾就是所谓的连锁店悖论(Chain-store paradox)。

针对连锁店悖论，克瑞普斯和威尔逊提出的思路是，[1]由于存在着信息不对称，垄断厂商发动价格战，从而使进入者认为自己"疯狂"，就有可能出现连锁店悖论。他们的模型假设在一个垄断市场，垄断者可区分为两种类型：其一是理性厂商，称作正常型(Normal type)厂商；其二是喜好不惜成本打价格战，而不喜欢和平共处的厂商，称作疯狂型(Crazy type)厂商。假设垄断厂商为疯狂型厂商的概率为 δ，而正常型厂商的概率为$(1-\delta)$，那么整个博弈的扩展式如图 5.6 所示。

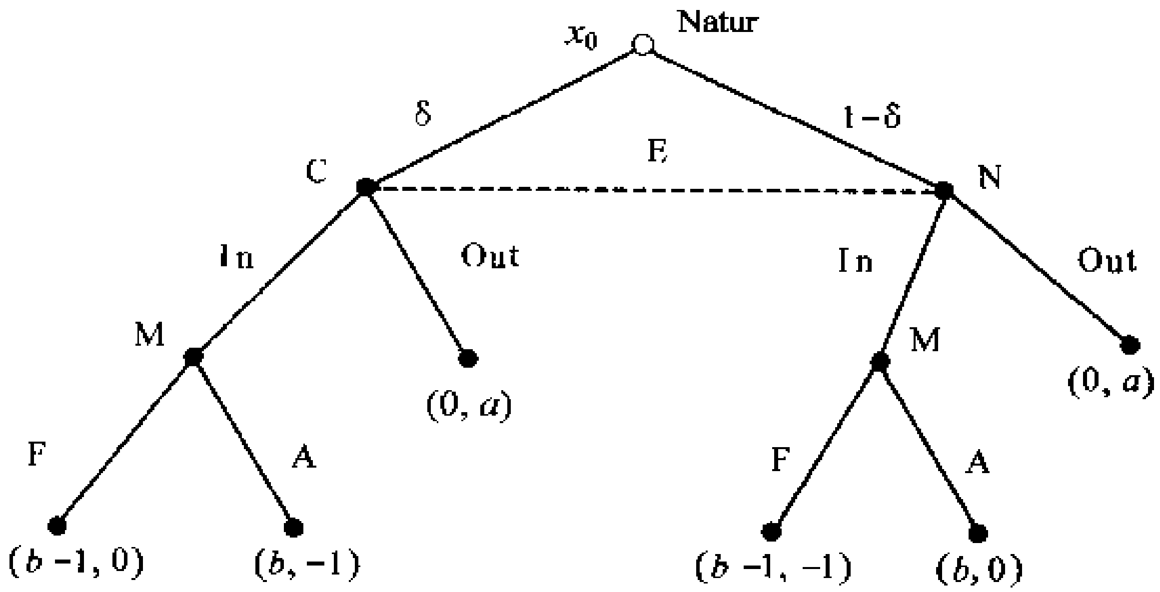


图 5.6 进入模型(2)

其中 C 表示疯狂，N 表示正常。图 5.6 的左半部分和右半部分惟一不同之处，在于垄断者选择 F(战斗)和 A(和平共处)的收益正好相反。很明显，疯狂型垄断厂商在遇到新进入者时，一定会进行价格战。在这个设定下，克瑞普斯和威尔逊对于价格战的解释如下：在进入博弈重复多期的情况下，[2] 由于进入者不知道垄断者的类型，因此正常型垄断

① Kreps, D. and Wilson, R. 1982:"Reputation and Imperfect Information". *Journal of Economic Theory* 27:253～279.

② 我们在介绍重复博弈时说过，阶段博弈不仅仅是静态博弈，他本身还可以是非完全信息动态博弈。

者有伪装成疯狂型厂商的动机。伪装的方法就是在新厂商进入时，即进行价格战。虽然这会受到损失，但如果因此使后来者误认为它是疯狂型而放弃进入，那么牺牲短期利润进行价格战以换取长期利润就是值得的。因此如果垄断者在市场上存在足够多期，那么在最初的几期里，任何类型的垄断厂商在遇到新进入者时，一定会发动价格战，因而没有厂商愿意进入。只有到最后几期，由于正常型厂商为了伪装所付出的代价已不足用长期利润（剩下的几期利润）来弥补时，才会暴露自己的真实类型。进入者也才会在最后几期真正知道垄断者的类型。只有它发现垄断者是正常型厂商时，才会选择进入。

假设有一个 N 期的垄断市场，垄断者 M 在这段期间内会留在市场中。M 的收益为每一期收益的总和。现在有 N 个潜在的进入者想进入垄断市场，以 $1,\cdots,n,\cdots,N-1,N$ 表示。也就是说，在第 n 期，厂商 n 和垄断厂商进行一次如图 5.6 所示的博弈。第 n 期，只有厂商 n 有进入垄断市场的选择。厂商 n 的收益即是它在当期进行该博弈的收益。令 p_n 为厂商 n 在进入前，认为 M 属于疯狂型厂商的概率。p_n 的值会随每期垄断者的行为而改变。因此，厂商 n 的策略，是根据它的猜测来决定要不要进入该市场。而 M 的策略，是在每一期，如果有厂商进入，它根据 p_n 的值来决定要不要进行价格战。

命题 5.1 以下的策略及猜测是进入博弈的序列均衡。

垄断者的策略：

M1. 假如垄断者属于疯狂型，则只要有新厂商进入该市场就进行价格战。

M2. 假如垄断者属于正常型，则在最后一期（$n=N$），垄断者会选择和新进入者和平共处。当 $n<N$ 时，如 $p_n\geq b^{N-n}$，其中 $1>b>0$，那么它一定进行价格战；如 $p_n<b^{N-n}$，那么它进行价格战的概率是 $(1-b^{N-n})p_n/(1-p_n)b^{N-n}$。换言之，如 $p_n<b^{N-n}$，那么垄断者使用混合策略$\{(1-b^{N-n})p_n/(1-p_n)b^{N-n}F,(b^{N-n}-p_n)/b^{N-n}(1-p_n)A\}$。

新厂商 n 的策略：

E1. 如果 $p_n>b^{N-n+1}$，则不进入市场（Out）。

E2. 如果 $p_n<b^{N-n+1}$，则进入市场（In）。

E3. 如果 $p_n=b^{N-n+1}$，则使用混合策略$(1/a\text{Out},(1-1/a)\text{In})$。

厂商对垄断者类型的猜测：

B1. $p_1=\delta$。

B2. 在任何 $N>n$ 的时期中，如厂商 $n-1$ 进入市场而垄断厂商采用和平共处策略 A，或 $p_{n-1}=0$，则 $p_n=0$。

B3. $n>1$，如 $p_{n-1}\neq 0$，而在第 $n-1$ 期时，新厂商 $n-1$ 进入市场而垄断者进行价格战，则 $p_n=\max(b^{N-n+1},p_{n-1})$。

证明：**B1** 不证自明，因为它是先验概率。**B2** 也很明显，因为如果垄断厂商选择和平共处，那么它一定是正常型厂商，所以 $p_n=0$。**B3** 的证明如下：如果 $p_{n-1}\geqslant b^{N-n+1}$，则根据垄断者的策略 **M2**，它应该在新厂商 $n-1$ 进入时，进行价格战。因此如果垄断厂商发动价格战，那么厂商 n 没有获得额外的信息，不会改变它对垄断者的猜测，因而 $p_n=p_{n-1}$。如果 $p_{n-1}<b^{N-n+1}$，那么根据 **M2** 垄断者应该在厂商 $n-1$ 进入时，用 $(1-b^{N-n+1})p_{n-1}/(1-p_{n-1})b^{N-n+1}$的概率进行价格战。因此，利用贝叶斯法则可得：

$p_n=$(M 是疯狂型的概率)/(M 打价格战的概率)

$=$Prob{M 是疯狂型}×Prob{M 进行价格战|M 是疯狂型}÷[Prob{M 是疯狂型}×Prob{M 进行价格战|M 是疯狂型}+Prob{M 是正常型}×Prob{M 进行价格战|M 是正常型}]

$$=\frac{p_{n-1}\times 1}{p_{n-1}\times 1+(1-p_{n-1})\dfrac{(1-b^{N-n+1})p_{n-1}}{(1-p_{n-1})b^{N-n+1}}}=b^{N-n+1}$$

因此 p_n 总是等于 p_{n-1} 和 b^{N-n+1}中较大的一个。

E1、**E2**、**E3** 的证明如下：

(1)若 $p_n\geqslant b^{N-n}$，则由 **M2** 可知，新厂商 n 进入后，无论是什么类型的垄断者都将进行价格战，因此新厂商选择不进入。

(2)若 $p_n<b^{N-n}$，则由于新厂商 n 进入后，正常型垄断者会用 $(1-b^{N-n})p_n/(1-p_n)b^{N-n}$的概率进行价格战，因此新厂商进入市场的

预期收益为：

$$p_n(b-1)+(1-p_n)\left[\frac{(1-b^{N-n})p_n}{(1-p_n)b^{N-n}}(b-1)+\frac{(b^{N-n}-p_n)}{(1-p_n)b^{N-n}}b\right]$$

$$=(b^{N-n+1}-p_n)/b^{N-n}$$

由于 $b^{N-n+1}<b^{N-n}$，因此 $p_n<b^{N-n}$ 时，E 应不应该进入要看 b^{N-n+1} 和 p_n 的值。如果 $p_n>b^{N-n+1}$，那么应该不进入；如果 $p_n<b^{N-n+1}$，那么应该选择进入。综合(1)、(2)可知，当 $p_n<b^{N-n+1}$，新厂商进入；当 $p_n>b^{N-n+1}$，新厂商不进入；当 $p_n=b^{N-n+1}$，则进入与不进入的预期收益相同，因此新厂商的策略为混合策略。为了计算的方便，不妨设以 $1/a$ 的概率选择不进入。

最后我们证明 **M1** 和 **M2**。**M1** 显然成立。我们只须证明 **M2**。若 $n=N$，即最后一期，正常型厂商和新进入的厂商将和平共处(已经没有战斗的必要)，收益为 0。如果 $n<N$，则令 $V_{up}(n)$ 是当 $p_n>b^{N-n+1}$ 时，正常型垄断者在第 n 期的预期利润；令 $V_{down}(n)$ 表示当 $p_n<b^{N-n+1}$ 时，垄断者在第 n 期的预期利润；令 $V(n)$ 代表 $p_n=b^{N-n+1}$ 时的预期利润。很明显，如果 $p_n>b^{N-n+1}$，那么新厂商不会进入(E1)，因此垄断者一直可以获得 a 的利润，而 p_n 的值会维持不变(因为没有厂商进入，因而后来的新厂商无法通过观察以前的历史来修正对垄断厂商类型的猜测)。但是 b^{N-n+1} 会随着时间的推移(n 变大)而增大。令 $m^*=\arg\min\{m|p_1\equiv\delta\leqslant b^{N-m}\}$，那么，在第 m^* 期之前的任何一期 $n<m^*$，由于 $p_n=p_1>b^{N-n}$，所以不会有新厂商进入(因为所有类型的垄断者都选择价格战)，垄断厂商会一直独享利润 a。到了第 m^*+1 期，新厂商 m^*+1 开始有可能进入，但在第 m^* 期，由于 $b^{N-m^*+1}<p_n\leqslant b^{N-m^*}$，垄断厂商采用混合策略或战斗，而新厂商不会进入，所以垄断者仍能获得 a 的利润。假如 $p_1=b^{N-m^*+1}$，垄断者的利润为 $V(m^*+1)$。若 $p_{m^*}<b^{N-m^*+1}$，那么垄断者的利润为 $V_{down}(m^*+1)$。因此：

$$V_{up}(n)=(m^*-n)a+a+V_{down}(m^*+1)\text{，其中 } n<m^* \tag{5.1.1}$$

$$V(n)=(1/a)(a+V_{down}(n+1))+(1-1/a)V_{down}(n+1)$$
$$=1+V_{down}(n+1) \quad (5.1.2)$$
$$V_{down}(n)=-1+V(n+1)\text{若 M 进行价格战或 0 若 M 选择和平共处。} \quad (5.1.3)$$

根据 **B3**，可知当垄断厂商进行价格战时，它在当期损失 1，但会使后面的新进入厂商修正猜测，即 $p_{n+1}=\max(b^{N-n},p_n)$，而 $p_n<b^{N-n+1}$新厂商才可能在第 n 期进入，所以 $p_{n+1}=b^{N-n}>b^{N-n+1}$，即(5.1.3)成立。

当第 n 厂商采取混合策略$(1/a\text{Out},(1-a)\text{In})$时，如果不进入，垄断厂商获得 a 的利润，一方面后面的新进入者对垄断厂商类型的猜测不会改变，即 $p_{n+1}=p_1=\delta=b^{N-n+1}$，但另一方面 b^{N-n+1}又会随 n 的增加而增大，即在 $n+1$ 期，$p_{n+1}<b^{N-n}$，即引起第 $n+1$ 厂商进入(**B2**)。如果第 n 厂商进入，那么垄断企业将采用混合策略(**M2**)，其预期利润等于0(因为和平共处的利润为 0，战斗的预期利润为 0，所以才采用混合策略)。但根据 **B3**，如果垄断厂商发动价格战，第 $n+1$ 厂商的猜测变为 $p_{n+1}=b^{N-n}$，但另一方面，如果垄断厂商采取和平共处，那么 $p_{n+1}=0<b^{N-n}$，综合两者可知，仍然会引起第 $n+1$ 厂商的进入。所以(5.1.2)成立。

利用迭代法，以及最后一期和平共处的结论，很容易解出 $V_{down}(N)=0$，从而 $V_{up}(n)=(m^*-n+1)a, n\leqslant m^*$；$V(n)=1$，而 $V_{down}(n)=0$。因此当一个垄断者在新厂商 n 进入时，如果它进行价格战，那么由于$p_n=\max(b^{N-n+1},p_{n-1})$，所以它接下来的利润为：

$$V(n+1)-1=0 \quad \text{若 } p_n\leqslant b^{N-n};$$
$$V_{up}(n+1)-1$$
$$=(m^*-n)a-1, m^*>n, a>1 \quad \text{若如 } p_n>b^{N-n}\text{。}$$[1]

而如果它不进行价格战，利润为 0。由此可知，当$p_n\leqslant b^{N-n}$，是否进行价格战对垄断厂商而言都一样(等于 0)，所以它用混合策略。而当

[1] 记住 $p_{n+1}=\max(b^{N-n},p_n>b^{N-n})$。

$p_n > b^{N-n}$时，进行价格战的预期利润大于0，因此它一定会进行价格战。

命题5.1表明，如果进入者相信垄断者可能会是不理性的(即使这个概率很小)，那么垄断者利用这个想法进行价格战以阻吓进入的做法，就是相当合理的。同时，如果期数很多，那么b^N会趋于0，从而小于$p_1=\delta$，因此，最初的几期新厂商都不会进入垄断市场。此外，由命题5.1可知，当$p_1=\delta$的值越大，m^*也越大，垄断厂商获得的垄断利润也越多。这就从理论上为连锁店悖论提供了一种合理的解释。

二、诉讼

诉讼博弈包含了博弈研究的几乎所有重要方面，例如信息不对称、信息揭露、谈判等，都可以在诉讼模型中自然地出现。诉讼中的原告和被告通常都拥有不同的私人信息，因而法律对原被告双方的举证责任有相当严格和细致的规定，而在诉讼过程中，双方提供的证据实际上就是一个信息揭露的过程，同时也是一个讨价还价的谈判过程。因而诉讼模型具有相当的代表性。这里我们考虑一个相当简化，但已包含上述所有特征的诉讼模型。

假设有一件人为意外的事件发生，受损者决定对惹事者提出赔偿要求。令P代表事件的原告，D代表事件的被告。被告所应承当的责任由q和w两个因素决定。其中w是被告对原告所造成的伤害，且原告和被告双方都知道w的值。q是被告应负担责任的比例。因此被告在这件事中，所应负担的赔偿金额为qw。我们假设q的值是被告的私人信息，诉讼的程序如下：原告先提出一个和解要求金额s，如果被告接受，则他付给原告s的赔偿金，这即是庭外和解。但如果被告拒绝这个提议，那么原告可以选择是否诉讼。假如原告认为自己胜诉的机会很低，那么他可以放弃诉讼。这时各自的收益为0。如果原告决定要上法庭解决，那么假设q的值会在法庭上被揭露出来。这时被告必须赔偿原告qw的金额，但是由于他们必须各自负担诉讼成本，所以原告实际上的收益为$(qw-C_q)$，而被告的收益为$-qw-C_d$，其中C_q和C_d是原告和

被告各自的诉讼成本。这个诉讼博弈可以用图 5.7 的博弈树表示出来。[1]

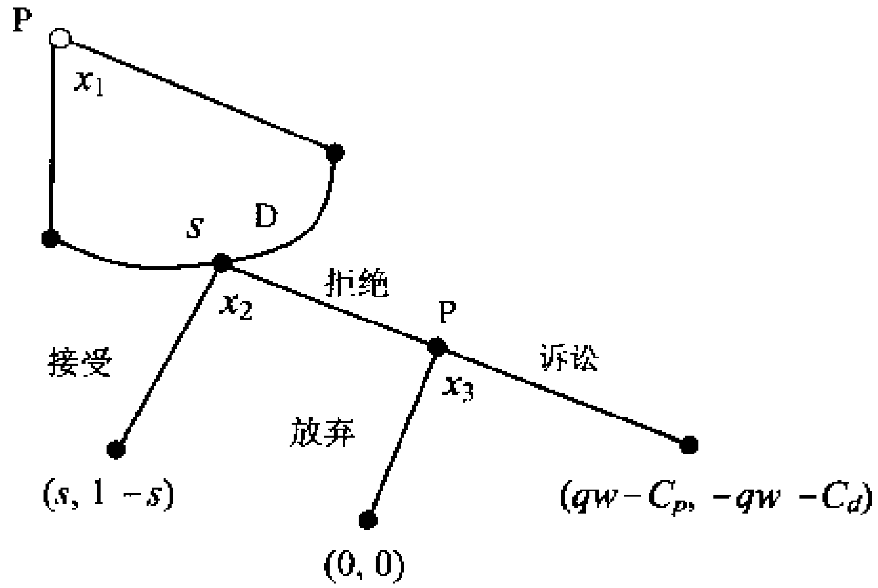


图 5.7　诉讼博弈

为了简化计算和分析有意义，我们假设：

A1. q 的值是在$[0,\bar{q}]$之间均匀分布，它的密度函数为 $f(q)$。

A2. 对原告来说，这个官司是值得打的，即 $E(q)w>C_p$。实际上这是一个先验态度。

我们用逆推法来求诉讼博弈的序列均衡。

第一步，先看节点 x_3 以下的子博弈，给定原告的提议 s 已经被被告拒绝，那么原告在决策点 x_3 就需要决定是否值得打官司，而这取决于他打官司的预期收益 $E(q|s$ 被拒绝$)w$ 和成本 C_p 的相对大小。若 $E(q|s$被拒绝$)w>C_p$，则官司值得打；若 $E(q|s$ 被拒绝$)w<C_p$，则不值得打这个官司。因此 P 的最优策略为：

如果 $E(q|s$ 被拒绝$)w>C_p$　则打官司；

如果 $E(q|s$ 被拒绝$)w=C_p$　则用 $v(s)$的概率打官司；

如果 $E(q|s$ 被拒绝$)w<C_p$　则放弃。

---

① 这个模型出自 Nalebuff, B. 1987:"Credible Pretrial Negotiation", *Rand Journal of Economics* 18:198～210。

但是E(q|s被拒绝)w的值是最初原告的提议s的函数。假如原告最初提议的s很小，那么如果被告还拒绝，表示q的真值一定非常小，否则被告不会冒上法庭的风险而去拒绝一个较低的和解要求，因为若上法庭，被告至少要支付C_d的诉讼成本。因此E(q|s被拒绝)的值必须往回推算，由被告在节点x_2的选择来决定。

第二步，被告在x_2是否接受提议s，主要取决于s与$r(s)(qw+C_d)$的相对大小，其中$r(s)$为一旦被告拒绝s后，原告打官司的概率。如果s非常小，那么被告应接受和解要求；如果s非常大，那么被告会拒绝。令$q(s)=[s/r(s)-C_d]/w$。$q(s)$是刚好令$s=r(s)(qw+C_d)$的q值，因此它是被告是否接受提议s的责任值q的临界点。如果$q>q(s)$，被告最优策略是接受s这个提议，如果$q<q(s)$，他应拒绝。因此我们可以用贝叶斯法则来计算，当s被拒绝时，原告对被告责任值的事后猜测为：

$$\mathrm{E}(q|s\text{ 被拒绝})=\int_0^{q(s)}\frac{qf(q)}{F(q(s))}\mathrm{d}q$$

现在再顺推回去。原告在x_3的选择，是计算$\int_0^{q(s)}\frac{qf(q)}{F(q(s))}\mathrm{d}q\times w-C_p$是否大于0。如果大于0，则预期利润$\int_0^{q(s)}\frac{qf(q)}{F(q(s))}\mathrm{d}q\times w$大于诉讼成本$C_p$，因此打官司是最优选择，否则放弃。由于$\int_0^{q(s)}\frac{qf(q)}{F(q(s))}\mathrm{d}q\times w$是$q(s)$的递增函数，因此$q(s)$越大，预期利润也就越大。因此我们可以找到一个$q(s)$的临界值，让原告打官司或不打官司的收益相同。令$q^*$是刚好让

$$\int_0^{q(s)}\frac{qf(q)}{F(q(s))}\mathrm{d}q\times w=C_p$$

成立的$q(s)$值。若$q(s)=q^*$，则原告接受s和打官司的收益完全相同；若原告对被告的q值的事后估计大于q^*，他会告下去，否则放弃，即如$q(s)>q^*$，则由于E(q|s被拒绝)$\times w>C_p$，因此原告在s被拒绝之后应打官司；如$q(s)<q^*$，E(q|s被拒绝)$\times w<C_p$，则原告在s被拒绝后会放弃；如$q(s)=q^*$，则原告随机地选择打官司。概括如下：

若 $q(s)>q^*$，则 $r(s)=1$；

若 $q(s)<q^*$，则 $r(s)=0$；　　　　(5.2)

若 $q(s)=q^*$，则 $r(s)\in[0,1]$。

由此我们可得到下列命题 5.2。

命题 5.2　当 $s>C_d$ 时，博弈在 x_3 点之下惟一的序列均衡是：

若 $s^*>C_d+q^*w$，则所有 q 值大于或等于 $(s^*-C_d)/w$ 的被告会接受 s^*，而所有 q 值小于 $(s^*-C_d)/w$ 则会拒绝，且如果 s^* 被拒绝，则 P 的最优策略是 $r(s^*)=1$。若 $s^*\leqslant C_d+q^*w$，则 q 值大于或等于 q^* 的被告会接受 s^*，而所有 q 值小于 q^* 则会拒绝，且 $r(s^*)=s^*/(q^*w+C_d)$。

命题 5.2 表明，如果和解提议 s^* 大于 C_d+q^*w，那么只有责任值 q 大于 $(s^*-C_d)/w$ 的被告会接受。而且，一旦 s^* 被拒绝，原告一定要告下去（正因为原告一定会告，所以被告不会拒绝）。如果和解提议 s^* 小于或等于 C_d+q^*w，那么只有大于 q^* 的被告会接受。而且一旦 s^* 被拒绝，原告会以 $s^*/(C_d+q^*w)$ 的概率告下去。

证明：如 $s^*>C_d+q^*w$，则由于 $r(s^*)$ 的值一定小于等于 1，因此 $q(s^*)=[s^*/r(s^*)-C_d]/w>q^*$。根据原告的最优反应(5.2)式，原告应让 $r(s^*)=1$。因此 $q(s^*)=[s^*/r(s^*)-C_d]/w=(s^*-C_d)/w$，由此可知 q 值大于或等于 $(s^*-C_d)/w$ 的被告会接受 s^*，而小于 $(s^*-C_d)/w$ 者会拒绝。

如 $s^*\leqslant C_d+q^*w$，且 $(s^*/r(s^*)-C_d)/w<q^*$，则根据(5.2)式，$r(s^*)=0$。但如此一来，上式左边不可能比右边小，因此不存在让上式成立的均衡。

如 $s^*\leqslant C_d+q^*w$，且 $(s^*/r(s^*)-C_d)/w=q^*$，那么我们由于 $r(s^*)\leqslant 1$，因此 $s^*\geqslant C_d+q^*w$。使上式成立的惟一办法，是调整 s^* 和 $r(s^*)$ 的值，让 $r(s^*)=s^*/(q^*w+C_d)$。而 $q(s^*)$ 的值则成为 $(s^*/r(s^*)-C_d)/w=q^*$，因此 q 值大于或等于 q^* 的被告会接受 s^*，而 q 值小于 q^* 者会拒绝。而在 s^* 被拒绝后，原告以 $r(s^*)=s^*/(q^*w+C_d)$ 的概率打官司。

在 $s\leqslant C_d$ 时，会有另外许多的均衡产生。比如说，原告提出一个小于 C_d 的提议 s，而被告接受。但万一被告拒绝，则由序列均衡在此对原

告的猜测没有限制，因此如果使用原告的猜测 $f(q)$，那么它将告下去。所以原告提出一个小于 C_d 的和解要求 s，被告同意；若被告拒绝 s，那么原告打官司这个策略组合，是一个序列均衡。但事实上原告要求 C_d 或比 C_d 小的和解金额 s，通常小得不合理。因此我们做一个比 A2 更强的假设：

A3. $E(q)w > C_p + C_d$。

在这个假设下，$s \leqslant C_d$ 的和解要求绝对不是原告的最优反应。因为 $s=C_d$ 比任何 $s<C_d$ 要好，而当 $s \leqslant C_d$ 时所有可能的被告都会接受。因被告无任何责任($q=0$)而被告上法庭的话，他必须支付 C_d，因而 C_d 是被告可能支付的最小值。如果 $s=qw+C_d$，那么事实上所有被告都会拒绝(因为这是被告可能支付的最大值)，而拒绝之后即使原告一定上法院，他可能获得的预期利润为 $E(q \mid s=\bar{q}w+C_d$ 被拒绝$)w-C_p=E(q)w-C_p$。但由 **A3** 可知，$E(q)w-C_p>C_d$，因此原告提议 $s=qw+C_d$ 比提 $s=C_d$ 要好。而提 $s=C_d$ 又比任何 $s<C_d$ 好，所以在 **A3** 的假设下，$s \leqslant C_d$ 绝对不会是原告的最优选择。所以我们可以完全不考虑 $s \leqslant C_d$ 的情形。接下来我们要在给定命题 5.2 的均衡策略下，求解原告在 x_1 的最优庭外和解要求 s^* 是多少。

第三步，令原告提议的和解要求为 s 时，他的预期收益为 $V(s)$，则：

$$V(s) = s[1 - F(q(s))] + (1 - r(s)) \times 0 + r(s)\int_0^{q(s)} (-Cp + qf(q)w)\mathrm{d}q$$

上式的第一项，代表被告接受和解时原告的收益。第二项代表被告拒绝和解，而原告决定不告下去的收益。第三项是被告拒绝和解，而原告继续告下去的预期收益。现在我们进一步假设：

A4. $\bar{q}w < 3C_p + C_d$。

那么可以证明：

(1)$V(s) < V(q^*w + C_d)$，$\forall\ s < q^*w + C_d$；

(2)$V(s) < 0$，$\forall\ s > q^*w + C_d$。

(1)的意思是说，如果原告提出的和解比 $q^{*}w+C_d$ 小的话，他的预期收益一定小于 $V(q^{*}w+C_d)$。换言之，均衡时原告的和解提议一定大于或等于 $q^{*}w+C_d$。而(2)的意思则是说，$V(s)$的值从 $q^{*}w+C_d$ 开始，是递减的。因此 $q^{*}w+C_d$ 是 $V(s)$在$(q^{*}w+C_d,\infty)$上的极大值。因此由(1)、(2)可以很容易知道，$V(q^{*}w+C_d)$是 V 的极大值。我们因此得到下面的命题。

命题 5.3　在 **A1**、**A3**、**A4** 的假设下，整个诉讼博弈惟一的均衡是：原告提出 $s^{*}=q^{*}w+C_d$ 的和解金额。如果被告的 q 值大于 q^{*}，则他接受提议 s^{*}，否则拒绝。如果 s^{*} 被拒绝，原告会用$r(s^{*})=s^{*}/(q^{*}w+C_d)=1$的概率打官司。

通过命题 5.3，我们可以预测一些类似于诉讼的模型均衡结果性质。比如，当被告或原告的诉讼成本越大时，原告所提出的和解金额会越大，而且原告所受到的损害值 w 与和解金额 s 无关。又由于和解概率是$(\bar{q}-q^{*})/\bar{q}$，因此由 q^{*} 的定义可知，当损害值越高或原告的诉讼成本越高时，和解的机率越高。

通过上面的两个例子我们可以看出，即使是非常简单的博弈，求解也非常不容易。而且在求解的过程中，通常还要加入一些假设，才能得到惟一的均衡结果。不过一旦求出均衡结果，通过举一反三，分析的威力就会体现出来。在阻吓性定价模型中，我们可以使模型进一步复杂化，例如考虑成本、市场需求等因素对均衡行为的影响，进而判断这个模型的现实解释力。在诉讼模型中，同样可以考虑诸如诉讼成本、损害大小、不确定的程度等因素对均衡的影响，以便了解诉讼制度本身的利弊。这种分析方式，即是近年来发展迅速的法经济学的一部分。它是利用经济分析的方法，来比较各种法律制度和规则，并研究这些制度的异同和效率性。实际上，很多法律制度和惯例，若从经济的角度来看，我们会很惊讶地发现，这些制度常常与经济原则不谋而合。比如契约法、诉讼法或破产法中的一些制度，若仔细分析，常常会发现它们的背后其实有效率性的考虑。而在商业法或民事诉讼法里，由于当事人双方的行为大部分就是经济行为，所以从经济学的角度来分析是相当恰当的。只是经济学里我们最注重的是效率，而通常没有注意到公平性。因此有时最

有效率的制度却不一定是公平的。但反过来说，如果一个制度没有效率，那么不论它是否公平，它都不是一个好的制度，因为我们这时可以用另外一个制度来增加某些人的福利，却不会因此而损害他人。所以分析一个制度的效率性，依然是一种判别制度优劣的方法之一，而博弈论就是分析制度形成与演变的有力工具。

第三节　信号博弈与序列均衡的再精炼

通过前面的介绍，我们已经看到序列均衡是非完全信息动态博弈的核心概念，所谓的非完全信息动态分析实际上就是求解序列均衡的过程，通过考察序列均衡的特点、性质、结构、条件，从而为各种非完全信息博弈提供解释。

在经济学中，信息不完全基本上又可等同于信息不对称，几乎所有的信息不完全模型都与信息不对称相关，逆向选择和道德风险就是信息不对称的两种表现形式。在这一节里我们将介绍信息不对称中的一个重要模型——信号博弈。信号博弈是研究逆向选择的一个重要模型，最早由斯彭思(Spence)提出并加以讨论，斯彭思也因此获得了2001年的诺贝尔经济学奖。

将序列均衡运用于实际分析面临着一个困难，就是有时求出的序列均衡存在许多个，这就给我们的进一步分析带来了相当大的困难，甚至无法获得更多的有用信息，使我们的分析失去意义。这表明序列均衡仍然存在缺陷，还需要进一步的提炼。我们将用信号模型来具体说明序列均衡的再精炼，这样做可以起到一石二鸟之效，既介绍了信号博弈，又说明了序列均衡的再精炼。

在介绍信号博弈之前，我们先讨论序列均衡存在的缺陷。

图5.8所示的博弈我们在前面已经反复用到，惟一不同的地方是对里面的收益值做了改动。根据序列均衡的定义，我们很容易找到图5.8所示博弈存在一个序列均衡，即{参与者1选择R，参与者2在信

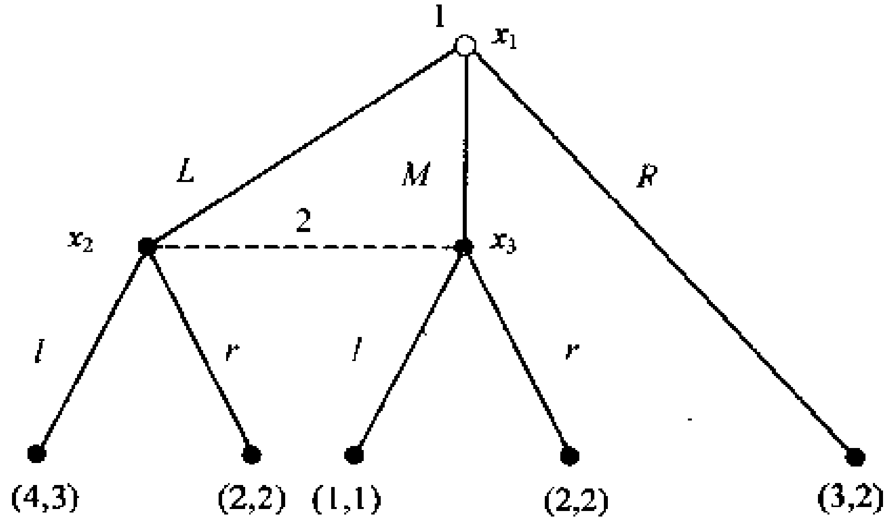


图 5.8 一个例子(4)

息集$\{x_2,x_3\}$选择r,$\mu_2(x_2)<1/2\}$。在猜测$\mu_2(x_2)<1/2$的情况下,如果参与者 2 选择l,那么预期收益为$3\mu_2(x_2)+1-\mu_2(x_2)=2\mu_2(x_2)+1$;如果选择$r$预期收益为$2\mu_2(x_2)+2(1-\mu_2(x_2))=2$。因而参与者 2 选择$r$就是最优反应。在给定参与者 2 策略的情况下,参与者 1 选择R的预期收益是 3,若他背离到M或L,收益为 2 或 3,因此参与者 1 选择R就是最优反应。

但是这组序列均衡却是不合理的。其原因是,如果参与者 1 选择R,则他的收益为 3;如果他选择M,则收益不是 1 就是 2。所以对参与者 1 来说,M是一个严格劣策略。早在第二章我们就已经证明过,无论在什么情况下,严格劣策略一定会被剔除,而不会成为均衡策略。因此假如博弈意外选择到信息集$\{x_2,x_3\}$,并让参与者 2 对信息集$\{x_2,x_3\}$作出猜测,那么他一定会认为$\mu_2(x_3)=0$。但是在上述的均衡里,参与者 2 却猜测,一旦博弈颤抖到信息集$\{x_2,x_3\}$,参与者 1 选择M而走到x_3的概率大于 1/2。,因而,上述序列均衡是一个非理性的均衡。之所以出现这种矛盾,在于序列均衡对于参与者的均衡外猜测(Out-of-equilibrium belief)仅要求策略是给定猜测下的最优反应,而没有进一步的限制。最重要的是,序列均衡完全漏掉了参与者会考虑他的对手为什么会背离原先的均衡这样的问题,因此常会产生一些不合理的均衡外猜测。在这一节里,我们主要的目的就是找出比序列均衡更合理,限制力更强

的设定均衡外猜测的方法。在信号博弈中，这些设定原则得到了充分的讨论和清楚的表述。

一、信号博弈的定义

最早的信号博弈模型是斯彭思 1973 年提出的求职模型。[①]在劳动市场上，通常企业主很难知道应聘者的能力高低，能力高低是应聘者的私人信息。因此企业主就需要一个可供判断的依据。斯彭思认为企业主对应聘者能力的高低判断，可以和能力有关，且需要从成本的信号中得出。这个信号愈强，成本愈高。而且能力愈高的人，愈容易发出强的信号。在劳动市场上，最常见的信号便是应聘者受教育的程度。因为求学的过程是要付出代价的，且能力愈强的人愈能忍受这个代价，也因此能够获得较高的学历。因而，学历高低在社会上就起着传递应聘者能力高低的信号的作用。也许文凭并不代表真正的能力，但文凭却有助于将高能力者从其他人中区分出来。劳动力市场模型是一个典型的信号博弈：

参与者 1(应聘者)拥有参与者 2(雇主)所不知道的私人信息(自身能力)，参与者 1 必须发出一个和自己私人信息相关的、需要成本的信号让参与者 2 接收。参与者 2 收到信号后，据此作出他的决策。确切地说，假设某个信号发送者(S)属于某种类型(type)t，且令 T 是所有可能类型的集合。发送者首先发出一个信号 $m(m\in M)$让接收者(R)接收。R 接收到 m 之后，采取回应 $a(a\in A)$。最后 S 和 R 所获得的收益是 t、m 和 a 的函数。

定义 5.6 一个信号博弈包含六个元素，记为 $\Gamma_S(N,T,M,A,\lambda,u)$。其中 $N=\{S,R\}$是参与者的集合。T 是 S 所有可能类型的集合。M 是所有可能的信号集合。A 是所有 R 的可能回应的集合[②]。λ 是信号接收者关

① Spence, A. M., 1973:"Job Market Signaling", *Journal of Economic Theory* 8:296～332.

② 在更复杂的信号博弈模型里，A 可以是 m 的函数。换言之，S 传出不同的讯号时，会使 R 有不同的回应集合。这时 $A=\{A(m)\mid m\in M\}$，不同的信号，会对应着不同的回应集合。我们不考虑这个情形。

于发送者类型的先验猜测。所以 $\lambda(t)\geqslant 0, \forall t\in T$，且 $\sum_{t\in T}\lambda(t)=1$。$u=(u_S,u_R)$，$u_i: T\times M\times A\to R, i=R,S$。即，每个参与者的收益是类型 t、信号 m 和回应 a 的函数。

在图 5.9 所示的博弈里，$T=\{t_1,t_2\}$，$M=\{l,r\}$，$A=\{u,d\}$。应聘者和雇主的预期收益分别为$u_S(t_1,l,d)=1$，$u_R(t_1,l,d)=2$，$u_S(t_1,r,d)=2$，$u_R(t_1,r,d)=3$，即第一个数为发信者的收益。

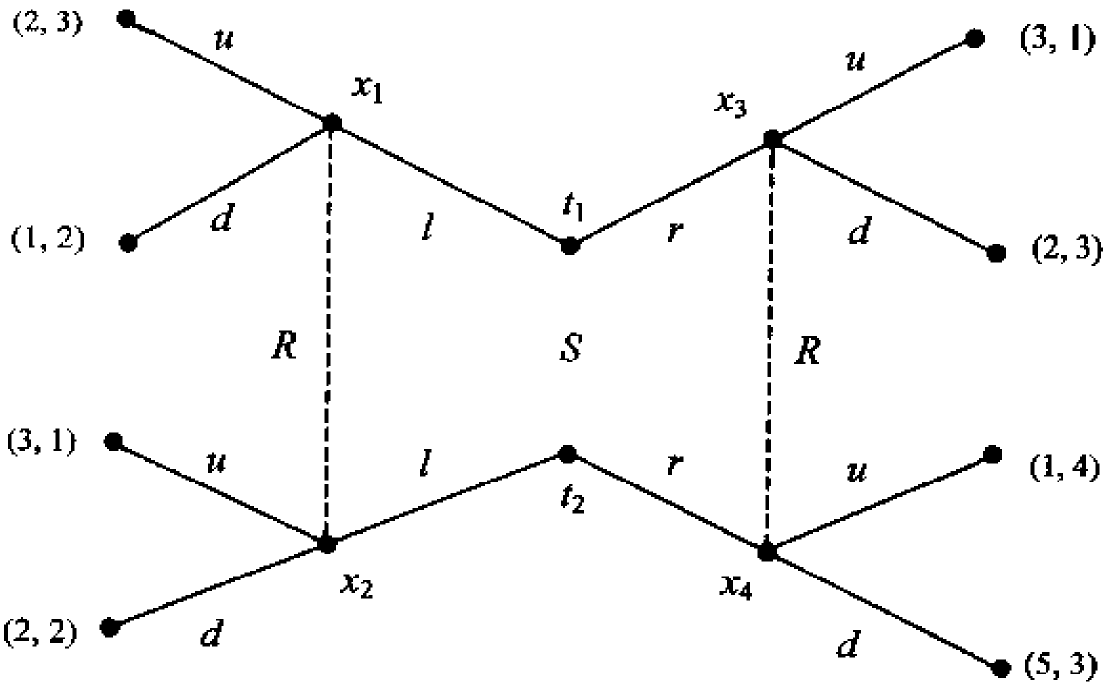


图 5.9　信号博弈(1)

定义 5.7　发信者的策略，是一个自他的类型映射到信号集合上的一个概率分布函数，$\sigma_S: T\to\Delta(M)$，其中 $\Delta(M)=\{(s_1,\cdots,s_{|M|})\in R^{|M|}\mid\sum_{i=1}^{|M|}s_i=1, s_i\geqslant 0\}$。收信者的策略是一个由信号映射到回应集合上的一个概率分布函数，即 $\sigma_R: M\to\Delta(A)$；其中 $\Delta(A)=\{a_1,\cdots,a_{|A|}\in R^{|A|}\mid\sum_{i=1}^{|A|}a_i=1, a_i\geqslant 0\}$。①

上述定义表明，信号发送者的策略是根据他自己的类型，来决定发

① $|M|$表示这个集合中所有元素的个数。

出什么样的信号。由于可以使用混合策略，因此他的策略是由 T 映射到信号集合(M)上的一个概率分布。同理，接收者的策略，是根据他收到的信号，决定如何回应。由于他也可以使用混合策略，所以是一个由 M 映射到 A 上的一个概率分布。

令 $\sigma_S(m;t)$是类型 t 的人，发出某个信号 m 的概率。而 $\sigma_R(a;m)$表示当收信者 R 收到信号 m 时，他回应某个 a 的概率。给定两个人的策略 σ_S 和 σ_R，一个类型为 t 的发信者，在(σ_S,σ_R)这组策略下的预期报酬是：

$$u_S(\sigma_S,\sigma_R,t) = \sum_{a\in A}\sum_{m\in M}u_S(t,m,a)\sigma_S(m;t)\sigma_R(a;m)$$

而接信者的预期报酬是：

$$u_R(\sigma_S,\sigma_R) = \sum_{a\in A}\sum_{m\in M}\sum_{t\in T}u_R(t,m,a)\sigma_S(m;t)\sigma_R(a;m)\lambda(t)$$

定义 5.8 如果一个信号博弈的一组策略和猜测($\sigma_S^*,\sigma_R^*,\mu^*$)符合下列的条件，那么它就是这个博弈的序列均衡。

(1)对于发信者的任一类型 $t\in T$，不等式

$$u_S(\sigma_S^*,\sigma_R^*;t)\geqslant u_S(\sigma_S,\sigma_R^*;t)$$

所有的 σ_S 成立，即任何一个类型发信者的策略，σ_S^* 是相对于 σ_R^* 和 μ^* 的最优反应。

(2)对于接信者的所有策略 σ_R，不等式

$$u_R(\sigma_S^*,\sigma_R^*,\mu^*)\geqslant u_R(\sigma_S^*,\sigma_R,\mu^*)$$

成立，即 σ_R^* 是相对于 σ_S^* 和 μ^* 的最优回应。

(3)$\mu^*(t|m)$代表当 R 收到 m 这个信号后，对 S 的类型的事后猜测(Posterior)。$\mu^*(t|m)$必须从 σ_R^* 和 σ_S^* 中通过贝叶斯法则得到。亦即对任意一个 $m\in M$，如果 $\sum_{t'\in T}\lambda(t')\sigma_S^*(m;t')>0$，则：

$$\mu^*(t|m) = \frac{\lambda(t)\sigma_S^*(m;t)}{\sum_{t'\in T}\lambda(t')\sigma_S^*(m;t')}$$

如果 $\sum_{t'\in T}\lambda(t')\sigma_S^*(m;t')=0$，则 $\mu^*(t|m)$ 可以任意给定。[①]

从上述定义可以看出，条件(3)与完美贝叶斯均衡的要求完全一样。在信号博弈中，完美贝叶斯均衡和序列均衡是完全等价的。这是因为在信号博弈中，对于均衡外信息集(Out-of-equilibrium information set)，我们可以(任意)用序列均衡的定义去无穷逼近原有策略，从而得到任何一种可能的猜测。因此信号博弈中的序列均衡对 R 在(意外)收到一个均衡外信号时，对 S 的类型所做的猜测，没有任何限制。这意味着在信号博弈中，可能存在着相当多的不合理的序列均衡，这也是为什么对序列均衡的再精炼会在信号博弈中得到讨论。

在信号博弈中，可能出现的情况是，不同类型的发信者都发出一样的信号，从而使得收信者无法对发信者的类型作出判断。在有些情况下，只有不同的类型对应着不同的信号，信号博弈才是有意义的。

定义 5.9　如果对于发信者的某个类型 $t'\in T$，对于任一类型 $\forall\, t\in T$，有 $\sigma_S^*(m;t')=\sigma_S^*(m;t)$，则 $(\sigma_S^*,\sigma_R^*,\mu^*)$ 叫做一个混同型(Pooling)均衡，相应的策略 σ_S^* 就称为混同策略；而如果对于某个类型 $t'\in T$，对于任一类型 $\forall\, t\in T$，有 $\sigma_S^*(m;t')\neq\sigma_S^*(m;t)$，则 $(\sigma_S^*,\sigma_R^*,\mu^*)$ 叫做一个分离型(Separating)均衡。相应的策略就称为分离策略。

混同型均衡是指不同类型的人，发出同一种信号。在混同型均衡下，收信者无法根据接收到的信号推断出发信者的类型。相反的，在分离型均衡下，不同类型的发信者会发出不同的信号，藉此，接收者可以完全推论出发信者的类型。这里我们着重讨论混同型均衡和分离型均衡，但并非所有的均衡不是混同型便是分离型均衡。有些时候均衡只能区分出部分的类型，亦即 $\sigma_S^*(t;m)\neq\sigma_S^*(t';m)$ 只对某些 $t,t'\in T$ 成立，我们把这种均衡定义为半混同型(Semi-pooling)均衡。如果某种类型 t 的发信者发出 m 的信号，而另一种类型 t^* 的发信者则随即地发出 m 和 m^* 的信号，这种均衡我们就称为杂合型(Hybrid)均衡。

考虑图 5.10 所示博弈。

---

① 这时由于 m 是一个不可能在均衡下被送出的信号，我们称 m 是一个均衡外信号(Out-equilibrium message)。

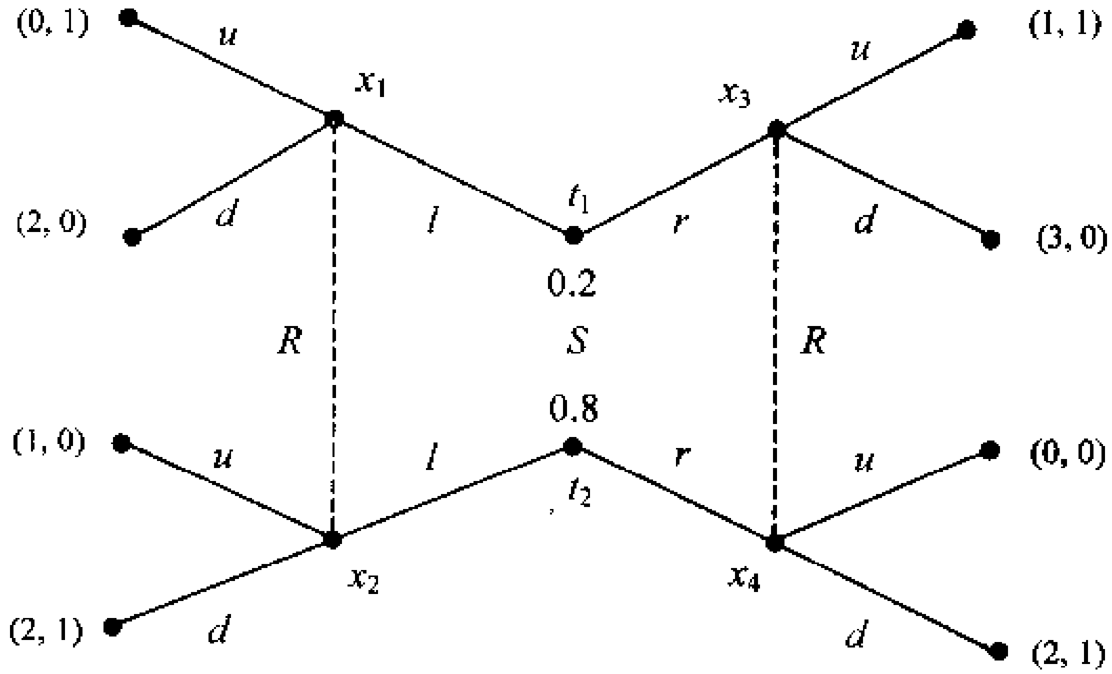


图 5.10 信号博弈(2)

其中 $\lambda(t_1)=0.2$，$\lambda(t_2)=0.8$。这个博弈存在很多的序列均衡，其中一个如下：$\sigma_S^*(r;t)=1\quad\forall t$，且 $\sigma_R^*(x_3,x_4)=d$；由于 $\{x_1,x_2\}$ 是一个均衡外信息集，不妨假设收信者的猜测 $\mu_R^*(x_1)=0.8$，$\mu_R^*(x_2)=0.2$，那么 $\sigma_R^*(x_1,x_2)=u$ 是 R 的最优反应。但这是一个很不合理的序列均衡。因为如果 S 是类型 t_1 的人，那么他选择 r 的收益一定会比选择 l 的收益大(因为 $3>2,1>0$)。如果 R 收到一个均衡外信息 l，那么他会认为 S 不太可能是属于类型 t_1 的人，或至少他应认为 S 属于类型 t_2 的可能性较大。但上面的序列均衡，非但如此，反而调高 S 是 t_1 的可能性猜测(由 $\lambda(t_1)=0.2$ 调整为 $\mu^*(t_1)=0.8$)。因此这是一个和直观相背的猜测设定。之所以如此，是因为序列均衡对均衡外猜测没有限制力。根据定义 5.8 中的条件(3)，当收信者接收到一个在策略均衡下不可能会收到的信息(Out-of-equilibrium message)时，他对发信者类型的猜测可以随意设定。

上述缺点表明，如果要在信号博弈中使得序列均衡更加合理，就必须在均衡外猜测上作出进一步的限制，使结果更合乎理性。以下我们所讨论的，是由赵(Cho)和克瑞普斯 1987 年提出的几个关于参与者如何

设定均衡外猜测的准则[1]。他们所使用的方法的基本精神是，当收信者收到一个原来均衡下，不该出现的均衡外信号 m 时，他应该考虑为什么传讯者会悖离原先的均衡，而送出 m 这个信号来。如果收信者对所收到的信号 m 作出各种可能的"合理"的回应，而在这些可能的回应之下，某种类型 t 的发信者所得到的收益，绝对不可能比原先的均衡收益高，那么发信者的类型就不应该是 t，否则他不会作出这种对自己有害无利的事情。因此，原先的序列均衡的猜测 μ^*，应该设定为 $\mu^*(t|m)=0$。如果原先的序列均衡没有这样设定猜测，那么这就是一个不合理的均衡，应当从均衡中剔除。所以这种做法，其实是在原先的序列均衡集合里找出在上述的精神下不合理的均衡，将之删除，而取剩下的均衡。但问题是，什么是收信者"合理"的回应？这其实没有标准答案。在不同的对所谓"合理"的解释下，我们会得到不同的均衡概念，以下我们将一一说明。

二、序列均衡的再提炼

所谓序列均衡的再提炼(Refinement)就是对已求解出来的均衡提出一些更强的要求和限制，从而达到强化均衡的目的。这些条件主要有以下三个：

1. 优势准则(Dominance criterion)

假设存在一个序列均衡$(\sigma_S^*, \sigma_R^*, \mu^*)$，在这个序列均衡之下，$R$ 收到的某个信息 m 的概率为 0，即 $\sum_{t\in T}\lambda(t)\sigma_S^*(m;t)=0$。如果存在 $t\in T$ 和 $m\in M, m'\neq m$，使得：

$$\min_{a\in A} u_S(t,m',a) > \max_{a\in A} u_S(t,m,a) \tag{5.3}$$

那么一个合理的均衡应该使 $\mu^*(t|m)=0$。如果一个序列均衡，符合上述限制，那它就满足优势准则。

这个准则非常直观：一个类型为 t 的人，如果他背离均衡，而发出

[1] Cho, I. K. and Kreps, D. 1987:"Signaling Games and Stable Equilibria", *Quarterly Journal of Economics* 102:179～222.

m 这个均衡外信号，那么他能期望得到最好的收益为 $\max\limits_{a\in A} u_S(t,m,a)$。如果这个类型的人有另外一个信号 m' 可供选择，而他发出 m' 这个信号后，所能得到的（可能）最差的收益 $\min\limits_{a\in A} u_S(t,m',a)$，都还比发送信号 m 所能得到的最大收益还高，那么发信者的类型不可能为 t，否则他不会做这种违反理性的事。因此如果(5.3)式成立，那么在原先的序列均衡下，在收信者收到 m 这个均衡外信号时，他的猜测 $\mu^*(t|m)$ 一定为 0，否则用优势准则的标准来看，原来均衡就是一个不合理的均衡，必须删除。

我们仍用图 5.8 中的序列均衡 $\sigma_1^*(x_1)=R,\sigma_2^*(x_2,x_3)=r,\mu(x_3)>1/2$ 为例。在 (σ_1^*,σ_2^*) 这个均衡下，M 和 L 都是均衡外信号。因为 $\min_{a\in A}u_1(t,R,a)=3$ 大于 $\max_{a\in A}u_1(t,M,a)=2$，所以根据优势准则 $\mu^*(1|M)=\mu^*(x_3)=0$。因此原先的均衡设定 $\mu^*(x_3)\geqslant 1/2$ 违反了优势准则的猜测，应该从均衡集合中删除。[①] 图 5.8 中的任何一个序列均衡，如果参与者 2 选择 R，那么满足优势准则的猜测 $\mu^*(x_3)$ 一定等于 0。但在 $\mu^*(x_3)=0$ 的情况下，$\sigma_2^*(x_2,x_3)=l$ 才是参与者 2 的最优回应。而给定参与者 2 的策略为 l，那么参与者 1 选择 L 就是最优反应。因而惟一一个序列均衡为：$\{\sigma_1^*(x_1)=L,\sigma_2^*(x_2,x_3)=l,\mu^*(x_3)=0\}$。这是惟一可以通过优势准则的序列均衡。

优势准则是筛选合理的序列均衡的方法之一，也是最粗浅的。由于它非常直观，所以它对均衡的限制力也很弱。比如在图 5.10 的均衡里，它就不能发挥作用：l 是一个均衡外讯号。现在 $\min_{a\in A}u_S(t_1,r,a)=1$，而 $\max_{a\in A}u_S(t_1,l,a)=2$。由于 $1<2$，所以(5.3)式不成立，我们因而也无法得到 $\mu^*(t_1|l)=0$ 的结论。

优势准则的问题是，它只是用发信者的某一信息 m' 下的最小收益和 m 下的最大收益相比，而完全没有考虑到收信者所做的回应 a，是否是相对于 m 和 m' 的最优回应，而不是故意去选一个 a 让发信者的收益最大或最小。这表明发信者在比较发送 m 或发送 m' 之后的收益时，仅

① 在这个例子里，参与者其实没有类型的分别，因此 $u_i(t,m,a)=u_i(m,a)$。

考虑(5.3)式,而没有进行更深入的比较。这就引出了所谓的直观准则。

2. 直观准则(Intuitive criterion)

在给出直观准则之前,我们先给出一些定义。令 $T(m)$ 为发信者所有可能发出 m 这个信号的类型集合,只有类型在 $T(m)$ 里的发信者,才有可能发出 m 这个信号。而 $\{\mu|\mu(T(m))=1\}$ 为在 $T(m)$ 上的收信者所有猜测的概率分布(密度函数)集合。设 $\{\sigma_S^*,\sigma_R^*,\mu^*\}$ 是一组序列均衡,而 m 是一个均衡外信号。令:

$$S(m)=\{t\in T|u_S^*(t)\geqslant \max u_S(t,m,a),a\in A_R(T,m)\} \quad (5.4)$$

其中 $u_S^*(t)$ 是类型为 t 的发信者在策略 $(\sigma_S^*,\sigma_R^*,\mu^*)$ 下得到的均衡收益,而

$$\begin{aligned} A_R(T,m) &= \bigcup_{\{\mu:\mu(T(m))=1\}} A_R(\mu,T(m),m) \\ &= \bigcup_{\{\mu:\mu(T(m))=1\}} \{a\in A|E_\mu u_R(t,m,a)\geqslant \\ &\quad E_\mu u_R(t,m,a'),\forall\ a'\in A\}。\end{aligned}$$

$A_R(T,m)$ 所代表的意思,是当收信者 R 收到 m 这个信号时,他会对发信者 S 的类型 t 有一个猜测 μ(这个猜测事实上是任意的①),在这个猜测 μ 下,R 就会有一个相对于 μ 的最优回应 a,而 $A_R(T,m)$ 这个集合即代表当 R 收到 m 这个信号时,在所有可能对 S 的类型的猜测下,R 的所有可能的最优回应集合。

(5.4)式可以解释如下:当收信者收到 m 这个均衡外信号时,他首先会建立一个对发信者类型的猜测:$\mu(t|m)$(在序列均衡下,这个猜测可以随意设定)。针对这个猜测,收信者采取相应的最优回应 a,即 $a\in A_R(T,m)$。假设发信者 S 的类型为 t,并发出信号 m,且收信者作出(在所有可能的猜测下)最优回应 $a\in A_R(T,m)$,那么 S 所有可能得到的收益中的最大收益为 $\max\limits_{a\in A_R(T,m)} u_S(t,m,a)$,如果这个最大收益也小于均衡收益 $u_S^*(t)$,那么(理性的)发信者绝对没有理由背离均衡,发出 m 这个信号。因为发信者背离均衡发出均衡外信号 m 所可能得到的最大收益仍然小于均衡收益。因此当接收者 R 收到 m 这个均衡外信号时,

① 在较一般的信号博弈里,猜测不会是任意的,因为有些类型的人根本无法送出 m 这个信号。

R 会认为发信者的类型一定不在 $S(m)$ 中，因为发信者偏离均衡所可能得到的最大收益小于遵循均衡得到的收益。由此我们看到，$S(m)$ 所代表的是不可能发送信号 m 的发信者的类型集合。

我们定义 $T|S(m)$ 为不包含 $S(m)$ 的所有类型集合。如果发信者发出了均衡外信号 m，收信者一定相信发信者的类型不在 $S(m)$ 中，如果在 $S(m)$ 中，发信者这样做就是非理性的，否则偏离到 m 显然毫无意义。但是，发信者却偏离了均衡，发出了均衡外信号 m，惟一的合理解释是：第一，发信者的类型不在 $S(m)$ 中；第二，发信者偏离均衡有利可图。这就意味着，存在某个类型 $t' \in T|S(m)$，使得：

$$u_S^*(t') \leqslant \min_{a \in A_R(T \cdot S(m), m)} u_S(t', m, a) \tag{5.5}$$

成立。只有在这种情况下，S 才有偏离的动机（更高的收益）和可能（S 不可能是类型 t 的人）。上述的论述显然十分的直观，因而被称为直观准则。

定义 5.10 直观准则：在信号博弈中，给定一组序列均衡 $(\sigma_S^*, \sigma_R^*, \mu^*)$，任何均衡外信号 m，如果满足 (5.4) 式和 (5.5) 式，那么一定有 $\mu^*(t|m)=0$。

直观标准具有一般性，不仅仅指信号博弈，只是信号博弈便于说明。实际上，任何一个合理的序列均衡，在任何均衡外策略 m 和类型 t 满足 (5.4) 和 (5.5) 两式时，应当有 $\mu^*(t|m)=0$。否则这个序列均衡就不符合直观准则，应该从均衡名单中删除。

接下来我们用图 5.10 来说明直观准则如何使用。

在图 5.10 所示的博弈中，我们给了一个在直观上不合理的序列均衡：$\{\sigma_S^*(r;t)=1, \forall\ t=t_1, t_2; \sigma_R^*(\{x_1, x_2\})=u, \sigma_R^*(\{x_3, x_4\})=d, \mu^*(x_1)=0.8\}$。我们现在要证明上述策略不满足直观准则。在上述的策略下，惟一的均衡外信号为 l，因此：

$$S(l)=\{t \in T \mid u_S^*(t) > \max_{a \in A_R(T, m)} u_S(t, l, a)\}=\{t_1\}^{①}$$

① $u_S^*(t_1)=3$ 大于 $\max\limits_{a \in A_R(T, m)} u_S(t_1, l, a)=2$，$u_S^*(t_2)=2$，等于 $\max\limits_{a \in A_R(T, m)} u_S(t_2, l, a)=2$。

由此可得 $T|S(l)=\{t_2\}$，$A_R(T|S(l),l)=d$，而 $\min\limits_{a\in A_R(T|S(l),l)} u_S(t_2,l,a)=2\geqslant u_S^*(t_2)=2$。因此，(5.4)和(5.5)两式皆成立。根据直观准则，$\mu^*(t_1|l)=0$，而这和上述策略下的猜测 $\mu^*(x_1)=0.8$ 相矛盾，所以上述策略虽是一个序列均衡，但不符合直观准则，因而应当从均衡集合中剔除。

三、重复直观准则(Iterated intuitive criterion)

对于有些博弈，运用一次直观准则并不能把所有的非理性(均衡外策略)序列均衡都剔除掉，而是需要反复运用直观准则。当(5.4)和(5.5)两式皆成立时，R 一定可以确定 S 的类型不在 $S(m)$ 里，所以 R 在做他的最优回应时，应该会将 $S(m)$ 从类型集合 T 中剔除。也就是说，这时候收信者根本不需要考虑 $S(m)$ 里的类型。换言之，在(5.4)式中，R 在做最优回应时，考虑所有的 $t\in T$ 显然不够聪明，他在选择最优回应时，事实上只需要考虑 $T|S(m)$ 里的元素就够了。因此，对于每个均衡外信号 m，定义 $T^2=T|S(m)$。令：

$$S^2(m)=\{t\in T^2|u_S^*(t)>\max_{a\in A_R(T^2,m)} u_S(t,m,a)\}^{①} \tag{5.6}$$

若存在 $t'\in T^2|S^2(m)$，使得：

$$u_S^*(t')\leqslant \min_{a\in A_R(T^3,m)} u_S(t',m,a) \tag{5.7}$$

那么根据直观准则，一个合理的均衡，必须令 $\mu^*(t|m)=0$，$\forall\, t\in S^2(m)$。也就是说，这时序列均衡的猜测，不应该认为发信者 S 属于$S^2(m)$里的类型。但如果满足(5.6)式和(5.7)式的 $S^2(m)$不是空集合，那么收信者在做最优回应时，应该把 $S^2(m)$从 $T|S(m)$中剔除。只要$S^k(m)$不为空集，那么这个程序就可以不断地重复进行。令 $T^k=T^{k-1}|S^{k-1}(m)$，$S^k(m)=\{t\in T^k|u_S^*(t)\geqslant \max\limits_{a\in A_R(T^k,m)} u_S(t,m,a)\}$。若存

① $S(m)$和 $S^2(m)$之不同处，在于前者考虑所有可能的类型，而后者只考虑在 T^2 里的类型。

在$t''\in T^k|S^{k-1}(m)$，使得$u_S^*(t'')\leqslant\min\limits_{a\in A_R(T^{k-1},m)}u_S(t'',m,a)$，则一个合理的序列均衡，必须令$\mu^*(t|m)=0,\forall t\in S^k(m)$。在经过$k$次重复直观准则的步骤后，收信者对于发出信号$m$的发信者的类型的猜测范围，就会缩小到$T^k|S^k(m)$。因为$T$是有限的，因而这个过程会在有限的次数内停止，而且最后剩下的$T^k|S^k(m)$不会是空集。上述这个过程被称为重复直观准则(Iterated Intuitive Criterion(IIC))。假使一个序列均衡的猜测μ^*，能够通过第k次的检验，那么我们称它为一个通过k次重复直观准则的均衡解。而如果一均衡猜测μ^*，能够通过所有层的重复考验(请注意这个过程在有限次内要停止)，那么就说它通过了重复直观准则。对重复直观准则的最形象的理解就是人与计算机下象棋的过程，在计算上显然人不如计算机来得快和准确，但人却能战胜计算机，原因就在于人们在计算如何走时，能够根据直观准则剔除明显不利的策略，从而使得自己考虑的策略集合大大缩小，而计算机却不具有像人一样的直观感觉，因而计算机相对于人而言，就是不聪明的，尽管它计算快、准确、能处理海量的数据。

四、斯彭思的就业市场信号模型

信号博弈的大规模深人研究始于斯彭思1973年提出的就业市场信号模型，该模型不仅开创了广泛运用扩展式博弈描述经济问题的先河，还较早地给出了完美贝叶斯均衡等概念。这里我们考察一个简单的就业市场信号模型。

假设有一个雇主(R)和求职者(S)。自然决定S有两种可能类型，高能力和低能力，$T=\{H,L\}$，高能力的先验概率$\mu(H)=1-q$。由于雇主不能直接观察到求职者的能力，因而只能通过一个与求职者能力相关的信号来对求职者的能力作出判断。这里我们不妨假设求职者发出的信号为他受教育的程度e，而雇主的最优回应是给出一个工资水平w。求职者受教育程度的高低主要取决于能力的高低，能力高的人显然获得同等的教育程度付出的成本要小于能力低的人，这是斯彭思就业市场模型的一个重要假定。即

$c_e(L,e)>c_e(H,e)$

其中 $c_e(t,e)$ 表示能力为 t、教育水平为 e 的求职者进一步接受教育的边际成本。根据这一假定，我们可知低能力的求职者的无差异曲线要比高能力的求职者的无差异曲线陡峭，可用图 5.11 表示出来。无差异曲线表示带给求职者不变的效用的工资 w 和受教育程度 e 的所有可能组合。受教育程度越高显然需要更高的工资来弥补付出的成本，所以无差异曲线斜率为正。由于低能力求职者要获得更高的教育水平需要付出更多的心血，于是就要求工资增加得更多些，才足以弥补他付出的努力。在图 5.11，教育水平由 e_1 增加到 e_2 时，低能力者要求工资从 w_1 增加到 w_L，而高能力者仅要求工资从 w_1 增加到 w_H。这里我们把受教育程度解释为学生在学校表现的差异，而不是受教育的时间年限。如果是受教育的年限，那么这个博弈会变成一个蜈蚣博弈，因为每一年求职者都会在是继续上学还是工作作出选择，如果选择工作，博弈结束，如果选择继续上学，那么博弈进行到下一阶段。

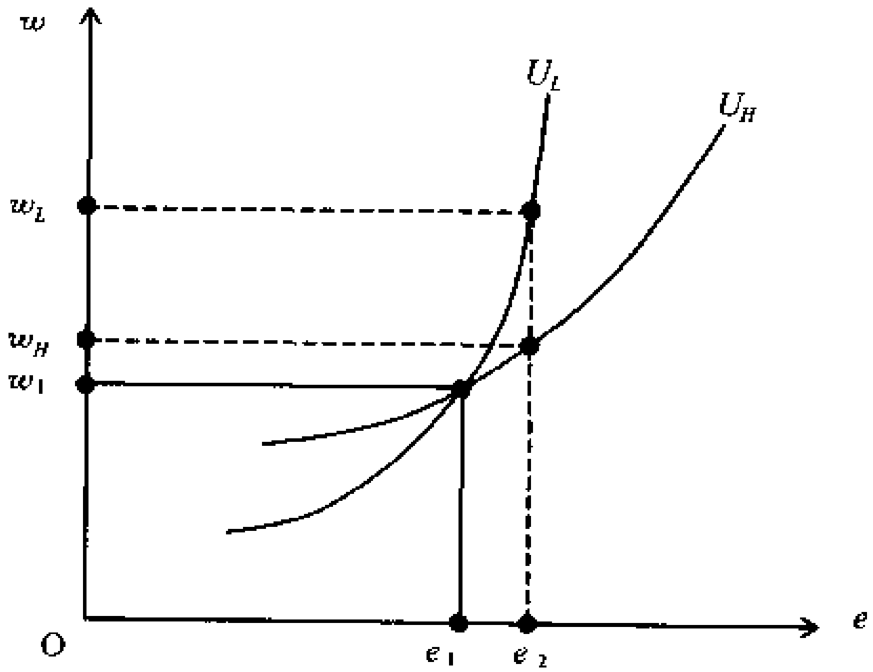


图 5.11　求职者的无差异曲线

为了简单，我们假设劳动市场处于完全竞争，因而雇主给出的工资水平为一个期望值，即 $w_p(e)=\mu(H|e)\times y(H,e)+[1-\mu(H|e)]\times$

$y(L,e)$。其中 $\mu(t|e)$表示雇主对求职者能力的猜测,而 $y(t,e)$则是求职者的产出。在这样的假设下,无论求职者的策略是什么,按期望产出支付工资都是雇主的最优策略,如果雇主相信求职者为高能力者,那么 $\mu(H|e)=1$,从而 $w=y(H,e)$。这样做有一个好处,就是我们可以专注研究信号者(求职者)的行为,而不用考虑收信者(雇主)的策略行为,同时又不失一般性。进一步的简化是假设产出 $y(t,e)$为线性(实际上这个假设可有可无)。那么在完全竞争情况下,求职者可得的工资 $w(e)=y(t,e)$,从而能力为 t 的求职者选择效用最大化的 e:

$$\max_e U_S(t)=y(t,e)-c(t,e)$$ ①

用 $e^*(t)$表示最优解,如图 5.12 所示,工资 $w^*(t)=y(t,e^*(t))$。

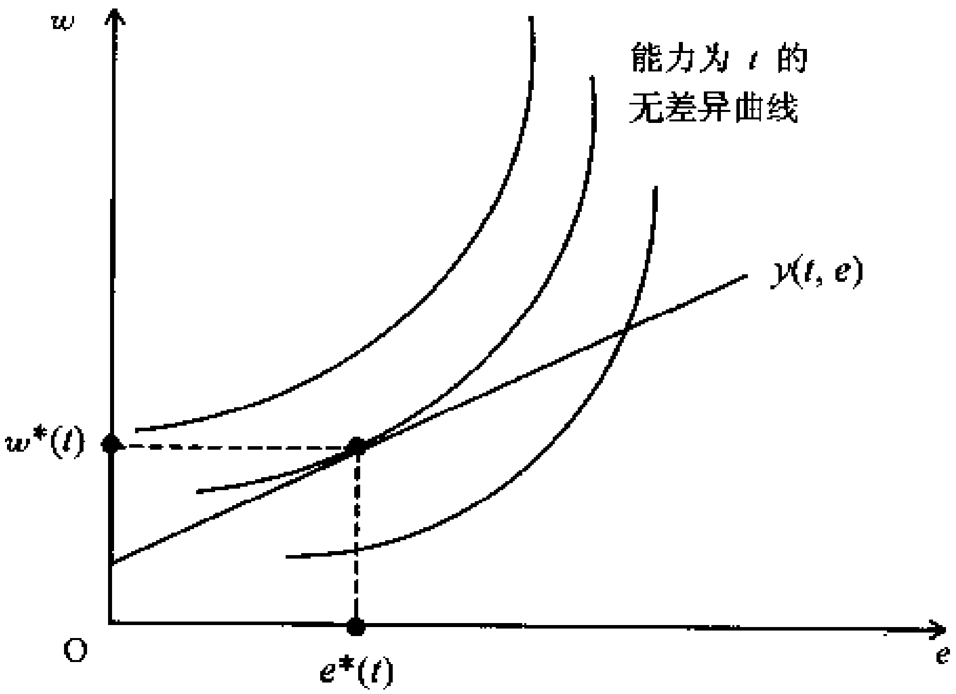


图 5.12 完全信息下的工资

在开始分析之前,有一个辅助定理可以帮助我们简化推论过程。

辅助定理 5.1 考虑上述就业市场模型中的某一个序列均衡,而 e 是这一均衡的均衡外信号。则不论收信者在这一均衡外信号 e 下的猜测是什么,这个均衡一定也可以被 $\mu'(L|e)=1$ 这个猜测所支持。也就

① 由于效用函数只考虑求职者,所以以后我们将下标 S 省略。

是说，在某一序列均衡下，在碰到任何一个均衡外信号 e 时，它对 S 的类型是 L（低能力）的猜测 $\mu^*(L|e)$，都可以用 $\mu^*(L|e)=1$ 去代替而使均衡行为不变。

辅助定理的证明，其实非常简单。证：一个序列均衡，在碰到某个均衡外信号 e 时，它对类型 L 的猜测 $\mu^*(L|e)=1-r$，不一定等于 1，但用 $\mu^*(L|e)=1$ 去替换之后，发出 e 的发信者，会被认为是类型 L 的求职者，从而得到低能力的人应得的收益，而这是最低的可能收益。即是说，任何一个求职者偏离到 e 时，在 $\mu^*(L|e)=1$ 这个猜测下，他会得到最低的可能工资。用公式表示即

$$\mu(H|e)y(H,e)+\mu(L|e)y(L,e)\geqslant 1\times y(L,e)$$

而在原先的猜测 $\mu^*(L|e)$ 下，原来的策略已是一个均衡，这说明求职者在原先的猜测 $\mu^*(L|e)$ 下，已经不愿意偏离到 e。现在如果用 $\mu^*(L|e)=1$ 去取代原有的猜测，那么偏离到 e 的求职者只可能得到更低的工资，因此他自然更不愿意偏离到 e。正因如此，用 $\mu^*(L|e)=1$ 去取代原有的任何一个序列均衡的猜测，不会改变原先的策略仍旧是一个均衡这样的结果。

利用辅助定理 5.1，我们可以很容易地找出就业市场模型的分离性和混同型均衡，更为重要的是，它可以大大简化我们的分析（只需要考虑 $\mu^*(L|e)=1$ 的情况，而不用考虑 $\mu^*(L|e)=1-r$）。令 $\hat{e}(t)=\arg\max_m y(t,e)-c(t,e)$。$\hat{e}(t)$ 是类型为 t 的求职者被认为他是类型 t 时，使他效用最大的信号。例如一个类型 L 的求职者，当别人也知道他的类型是 L 时，令他的效用最大化的信号即是 $\hat{e}(L)$。令 $\underline{U}(L)=U_S(L,\hat{e}(L),y(L,e))$。很明显，一个类型 L 的求职者，在均衡时他最低的效用绝不会低于 $\underline{U}(L)$（因为最差不过讲真话）。令 $\hat{e}(H,L)=\arg\max_m y(L,e)-c(H,e)$。$\hat{e}(H,L)$ 是类型 H 的求职者被人误认为他是类型 L 时，使他的效用达到最大化的信号，这时的效用定义为 $\underline{U}(H)=U_S(H,\hat{e}(H,L),y(L,e))$。$\underline{U}(H)$ 代表一个类型 H 的求职者，在均衡时所能得到的最低效用。以上叙述可以绘成图 5.13。图中越往上的无异曲线效用越高，同时类型 L 的无异曲线的斜率，也比类型

H 者陡峭。

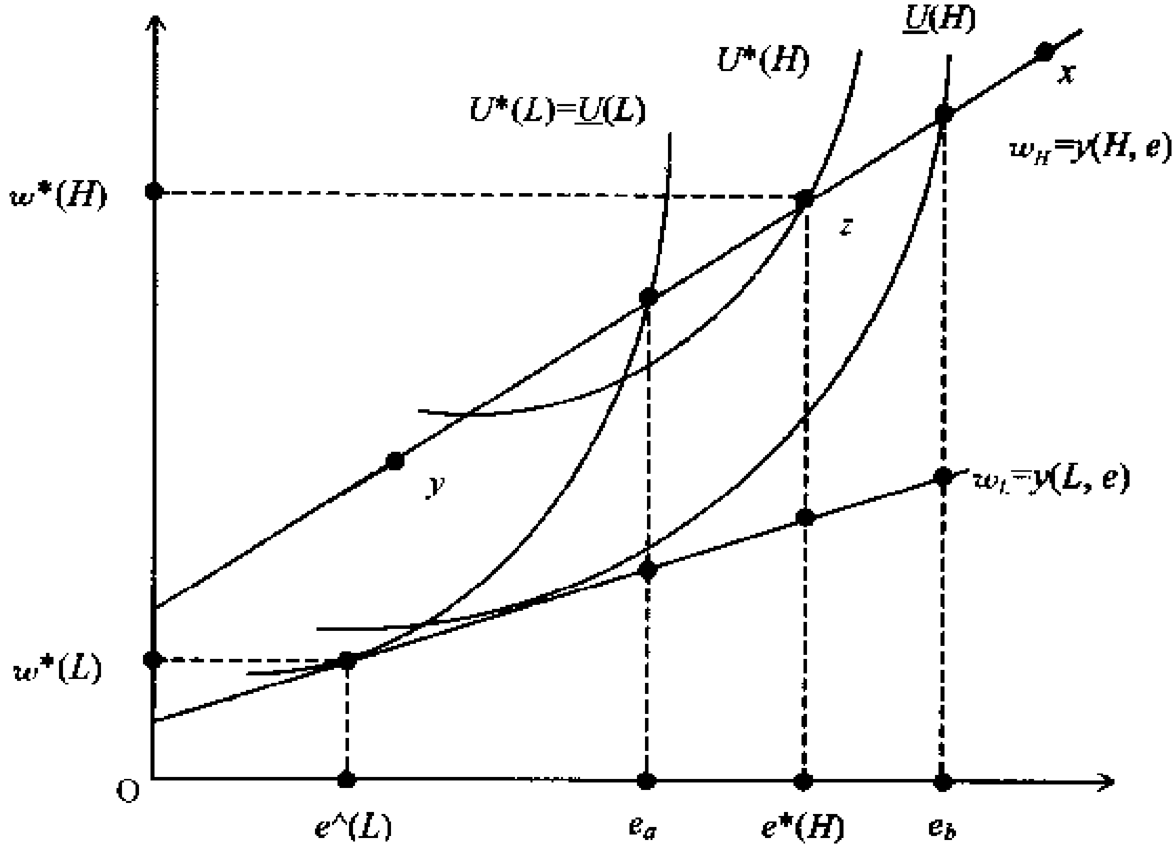


图 5.13 分离型均衡

1. 分离型均衡

在分离型均衡下，两种类型的求职者会发出不同的教育水平信号，因此在均衡状态下，雇主可从求职者发出的信号来判断他属于何种类型。这表明类型 L 的求职者的工资会在 $w_L=y(L,e)$ 线上的某一点，而类型 H 的人的工资会在 $w_H=y(H,e)$ 线上的某一点。由于类型 L 的求职者是生产力最低者，所以他一定会选择一个 e 让自己的效用最大，而这个信号其实就是 $e^{\wedge}(L)$。即，类型 L 的求职者的无异曲线会和 $w_L=y(L,e)$ 这条线相切于 $e^{\wedge}(L)$，而他所得的效用正好为 $\underline{U}(L)$。因此，在分离型均衡之下，$e^*(L)=e^{\wedge}(L)$，而 $\underline{U}(L)=U^*(L)$。接下来我们看类型 H 的求职者。

他所发出的均衡信号，$e^*(H)$，不会大于 e_b。这是因为如果 $e^*(H)>e_b$，那么类型 H 的求职者得到的工资会在 $w_H=y(H,e)$ 线上，

e_b 之右的某一点上(例如图 5.13 中的 x 点)。但如果这样,他宁可发出 $\hat{e}(H,L)$这个信号。这是因为如果他发出 $\hat{e}(H,L)$,那么根据辅助定理5.1,他会被认为是类型 L 的求职者,得到 $w=y(L,e)$的工资,因此得到$\underline{U}(H)$的效用。而由图 5.13 中可以看出,通过 x 点的无异曲线的效用要小于$\underline{U}(H)$。因此偏离到 $\hat{e}(H,L)$所得到的效用会比发出 $e^*(H)$的效用要高。

$e^*(H)$也不会小于 e_a。因为如果 $e^*(H)<e_a$,那么他会得到像图 5.13中 y 点的工资。但在 y 点,类型 L 的求职者所得到的效用大于 $U^*(L)$,因此他一定会偏离 $e^*(L)$,而发出 $e^*(H)$这个信号,这样他会被认为是类型 H 的求职者,从而得到更高的效用。这和 $e^*(L)$是均衡信号的结论相矛盾。因此 $e_a \leqslant e^*(H) \leqslant e_b$。

命题 5.4　只要任何一个 $e^*(H)$满足 $e_a \leqslant e^*(H) \leqslant e_b$,那么下列策略和猜测是分离型均衡。

(1)$e^*(L)=\hat{e}(L)$,$e^*(H)=\hat{e}(H)$;

(2)$\mu^*(L|e^*(L))=1$,$\mu^*(H|e^*(H))=1$,$\mu^*(L|e)=1$,

$\forall e \neq e^*(L),e^*(H)$。

证明:由于分离型均衡的特性,再加上辅助定理 5.1 的结论,命题 5.4 中的条件(2),对 μ^* 的设定明显成立。因此我们只要证明在 μ^* 的设定下,$e^*(L)$和 $e^*(H)$各自是求职者的最优回应即可。如果类型 L 的求职者没有发出 $e(L)=\hat{e}(L)$,而发出 $e^*(H)$,那么他会被认为是类型 H 的求职者,而得到的工资会在$w_H=y(H,e)$这条线上,且对应于 $e^*(H)$的 z 点上。但由于 $e^* \geqslant e_a$,因此由图 5.13 可以很容易看出,类型 L 的求职者在这点所得到的效用反而变低了。因此类型 L 的求职者不会发出 $e^*(H)$的信号。如果他发出任何不等于 $e^*(H)$的讯号,那么由 μ^* 的设定,他会被认为是类型 L 的求职者。但由 $\hat{e}(L)$的定义可知,$\hat{e}(L)$已经是类型 L 的最优信号。因此类型 L 的求职者发出 $\hat{e}(L)$是他的最优策略。接着我们要证明 $e^*(H)$是类型 H 的求职者的最优策略。如果他偏离到任何一个不等于 $e^*(H)$的信号,那么由 μ^* 的设定,他会被认为是类型 L 的求职者,因而所得到的工资会在 $w_L=y(L,e)$这条线上。由图 5.13 可以看出,这样的效用绝对不可能比

$U^*(H)$大。这就证明了命题 5.4。

2. 混合型均衡

在混合型均衡下，类型 L 的求职者和类型 H 的求职者会发出一样的信号，因而得到一样的工资。令 $e^*(L)=e^*(H)=e^*$。由于两个求职者的信号相同，所以雇主对求职者的猜测仍是原先的先验猜测，也就是 $\mu(L)=q$。由于求职者的工资是雇主对他生产力的预期，因此在均衡状态下，两个求职者的工资在 $w_p=\mu(L)y(L)+\mu(H)y(H)$ 这条线上（见图 5.14）。

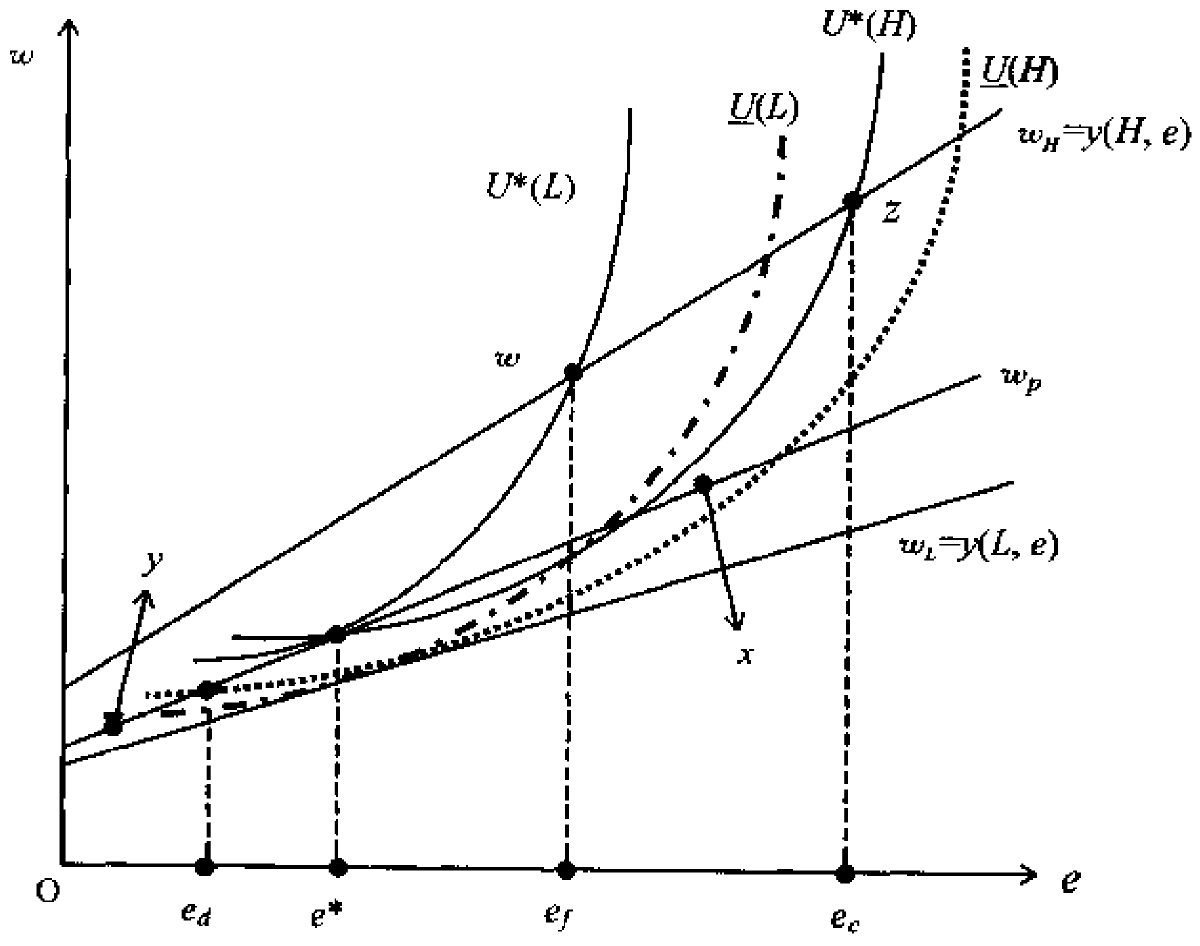


图 5.14 混同型均衡

我们要证明，在均衡的时候，两个求职者所发出的信号，一定要在 e_d 和 e_f 之间。这个原因很简单。如果 $e^*>e_f$，那么类型 L 的发信者所得到的效用，会比 $U^*(L)$还要低。比如，无差异曲线经过图 5.14 中的 x 点的类型 L 的求职者，所得到的效用会在 $U^*(L)$这条无异曲线之下，

因此类型 L 的求职者愿意偏离 e^* 而发出 $e^{\wedge}(L)$ 的信号，以得到 $U^*(L)$ 的效用。如果 $e^*<e_d$，那么类型 H 的发信者所得到的效用，会比 $\underline{U}(H)$ 还要低。例如，经过图 5.14 中 y 点的类型 H 的求职者，所得到的效用会在 $\underline{U}(H)$ 这条无异曲线之下，因此类型 H 的求职者宁可偏离 $e^*(H)$ 而送出 $e^{\wedge}(H,L)$ 的信号，以得到 $\underline{U}(H)$ 的效用。不但如此，我们还可以证明在 e_d 和 e_f 之间的任何信号 e^*，都可以是混同型均衡。

命题 5.5　只要 $e_f\geqslant e^*\geqslant e_d$，那么下列策略和猜测是混合型均衡。

(1) $e^*(L)=e^*(H)=e^*$；

(2) $\mu^*(L|e^*)=q,\mu^*(H|e^*)=1-q,\mu(L|e)=1,\forall\ e\neq e^*$。

证明：(2)的证明很容易，请读者自己证明。证明(1)等价于证明 $e^*(L)$ 和 $e^*(H)$ 各是对方的最优回应。定义 $\max U^*(t)=w_p-c(t,e)$。如果类型 L 的求职者发出任何一个 $e\neq e^*$ 的信号，那么由(2)的设定可知他会被认为是类型 L 的求职者。但由于 $e_f\geqslant e^*$，因此很容易可以由图 5.14 得出 $U^*(L)>\underline{U}(L)$。由此可知 $e^*(L)$ 是类型 L 的求职者的最优回应。如果类型 H 的求职者发出任何一个的 $e\neq e^*$ 信号，那么由(2)可知，他会被认为是类型 L 的求职者。因此他最好也只能得到 $\underline{U}(H)$ 的效用。但由于 $e^*\geqslant e_d$，因此由图 5.14 很容易看出 $U^*(H)>\underline{U}(H)$。

命题 5.4 和命题 5.5 所找到的均衡其实并不是所有的均衡，因为还有很多均衡既不是混同型也不是分离型，但那些均衡一方面求解不易，另一方面也不太有直觉上的意义，因此我们只对上述两类的均衡有兴趣。接下来我们要看看，是不是所有的上述的均衡，都可以通过直观准则。

先看混同型均衡。类型 L 的求职者，如果发出一个在 e_f 和 e_c 之间的信号的话，得到的工资最多也只会在 $w_H=y(H,e)$ 这条线上的 w 和 z 之间的点上(见图 5.14)。但类型 L 的求职者在这里任何一点的效用，都低于 $U^*(L)$。因此对于任何 $e'\in(e_f,e_c)$，$L\in S(e')=\{t\in T\mid U^*(t)\geqslant\max\limits_{w\in A_R(T,e')}U_S(t,e',w)\}$。这表明，类型 L 的求职者根本不会送出 e' 这个信号。因此，如果雇主收到 e'，他应会认为求职者是类型 H 的求职者，而给他 $\overline{wz}$ 这条线段上对应于 e' 的工资。一旦如此，类型 H 的求

职者的效用会大于均衡效用$U^*(H)$。所以类型H的求职者一定有动机偏离到e'，以证明自己是类型H的求职者，从而得到更高的效用。$\min\limits_{w\in A(T,L,e')} U^*(H,e',w) \geq U^*(H)$，因此(5.4)和(5.5)式成立。根据直观准则，我们有$\mu^*(L|e')=0$。因此原先的所有混同型均衡中的猜测，$\mu^*(L|e')$必须等于0。一旦如此，类型H的求职者一定会偏离原先的均衡信号e^*，而发出$e'\in(e_f,e_t)$这个信号。所以类型H的求职者一定可以找到一个$e'\in(e_f,e_t)$使得原先的均衡不能满足直观准则，由于e^*可以为e_d和e_f之间的任意数，所以没有一个混合均衡能够通过直观准则。

接下来我们再看看分离型均衡是否可以通过直观准则。令$e\in(e_a,e_b)$。一个类型L的求职者，如果发出e这个信号，那么即使他得到类型H的求职者的工资，他的效用也会低于$\underline{U}(L)$(见图5.13)。因此$\underline{U}(L) > \max\limits_{w\in A_R(T,e)} U^*(L,e,w)$。这表明$S(e)=\{L\}$。但由于$T|S(e)=\{H\}$，且由图5.13可得$U^*(H) < \min\limits_{w\in A_R(T|S(e),w)} U^*(H,e,w)$。因此由直观准则可知$\mu^*(L|e)=0$。这意味着，类型$L$的求职者如果发出$e$这个均衡外信号，一定会使自己的效用水平变低，因此雇主会认为类型L的求职者不会发出e这个信号，即$S(e)=\{L\}$。但如果雇主真这么认为，那么在收到e后就会给求职者类型H的高工资，那么类型H的求职者就愿意偏离到$e(<e^*(H))$这个均衡外信号，从而使新的效用高于原来的均衡效用$U^*(H)$。所以直观准则要求当雇主收到e这个信号时，应该把他对求职者是属于L的猜测调为0，亦即$\mu^*(L|e)=0$。换言之，在原先所有的分离型均衡里，我们必须让$\mu^*(L|e)=0$。但如果这样，类型H的求职者一定会偏离$e^*(H)$而发出$e\in(e_a,e^*(H))$的信号，因此原先的均衡就不再是均衡。这个情形的产生，完全是因为e_a和$e^*(H)$之间的信号具有上述特性。所以要使类型H的求职者不会偏离$e^*(H)$，惟一的可能是令$e^*(H)=e_a$。所以$e^*(H)=e_a$是分离均衡下，类型H的求职者惟一可能发出的信号。

综上所述可知，在就业市场信号模型中，虽然存在着许许多多的分离型、混合型、半混同型、杂合型的均衡，但除了惟一一个分离型均衡能

通过直观准则的检验外，其余的序列均衡没有一个能满足直观准则。直观准则的使用，剔除了几乎所有的序列均衡，而只留下对均衡的惟一一个预测。

第四节　结束语

人类社会的一切矛盾实质上都根源于人们之间的经济利益冲突，人类社会的各种制度的建立，大到上层建筑，小到单位的规章，实际上都是为了解决人类社会的各种冲突而作出的一种选择。从经济学本身来看，脱离社会关系的经济学必然是残缺的和无意义的，然而要克服经济学的这种残缺除了勇气和毅力，还需要达成目标的手段，而博弈论就是这样一种手段，它使经济学变得丰满了起来，使经济学鲜活了起来，也使人类理解自身的社会性有了深人的可能。亚里斯多德早在几千年前就曾说过，"人是社会性动物"，而博弈论就是我们打开人的社会性大门的钥匙。也许这是对学习博弈论有什么用的最好的回答。

在本书的最后，我们对所学习的博弈均衡概念作一个简单的总结。从博弈论中的均衡概念可以看出，序列均衡是最一般的均衡概念，它包含着所有其他均衡概念，子博弈完美纳什均衡、贝叶斯纳什均衡和纳什均衡都是序列均衡分别在动态完全信息博弈、静态非完全信息博弈和完全信息静态博弈下的特例。一个序列均衡一定是子博弈完美纳什均衡、贝叶斯纳什均衡和纳什均衡，但反之不成立。从纳什均衡向上走，每一个均衡概念限制的条件会越来越强，子博弈完美均衡是一个剔除不可置信威胁的纳什均衡，而贝叶斯纳什均衡则是在理性猜测下的纳什均衡，到了序列均衡不仅要求剔除不可置信威胁，而且要求更强的猜测设定，所有策略都是在给定猜测下的最优策略。从理论上讲，运用序列均衡概念几乎可以分析所有的社会经济现象，但令人遗憾的是，至今求解序列均衡仍没有一般可用的方法，哪怕一个简单的模型求解序列均衡也十分繁琐，通常需要我们先去猜出可能的均衡，然后再加以证明。

尽管如此，博弈论仍然为经济学的发展开辟了一场真正意义上的革命。通过博弈论提供的理论，制度的演化成了经济学研究的对象。通过人工智能模拟各种可能的演化机制，为法经济学、行为经济学与实验经济学等的研究提供了基本的条件，而这种研究的新方向构成了今天经济学最为激动人心的领域。

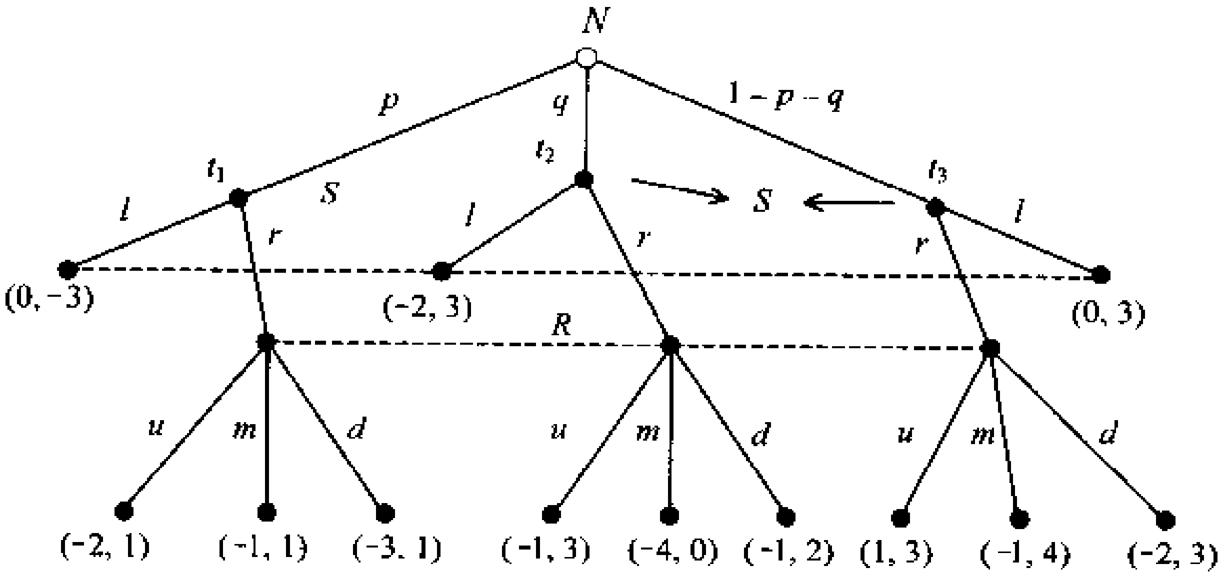
总而言之，掌握博弈论已成为从事经济学研究必备的基本素质。

复习思考题

1. 什么是完美贝叶斯均衡、序列均衡、直观准则？完美贝叶斯均衡存在什么缺陷？序列均衡存在什么缺陷？序列均衡和直观准则的意义是什么？

2. 请你查到 Nalebuff(1987 年)的论述诉讼的文献，并且读懂它，同时细心体会如何求序列均衡。

3. 请你找出下图所示博弈中的所有序列均衡。解释你的结论，并用优势准则、直观准则和重复直观准则来检查你的答案。



一个信号博弈

参考书目

1. Gibbons, R. 1992: *Game Theory of Applied Economists*. Princeton: Princeton University Press.
2. Kreps, D. M. 1990: *A Course in Microeconomic Theory*. Princeton: Princeton University Press; Part III.
3. Mas-Colell, A., Whinston, M. and Green, J. 1995: *Microeconomic Theory*. Oxford: Oxford University Press. Chapter 6-8.
4. [美]罗伯特·吉本斯. 博弈论基础. 北京:中国社会科学出版社, 1999.
5. [加]马丁·J. 奥斯本,[美]阿里尔·鲁宾斯坦. 博弈论教程. 北京:中国社会科学出版社,2000.
6. [美]迈尔森. 博弈论——矛盾冲突分析. 北京:中国经济出版社, 2001.
7. 张维迎. 博弈论与信息经济学. 上海:上海三联书店、上海人民出版社,1996.

A1108124

后 记

在我刚开始上研究生的时候，由于一次非常偶然的课堂讨论，使我对一个理论问题百思不得其解，这是一个关于显示偏好的问题。由于教材上只讲授了显示偏好弱公理和显示偏好强公理，而没有介绍显示偏好一般公理，使我当时以为自己找到了一个发展重大理论的机会，后来才知道只不过是自己的一厢情愿。但是，通过对显示偏好理论的学习和“研究”，我开始接触到了机制设计理论（实际上就是信息经济学），通过机制设计理论又接触到了各种各样的逆向选择和道德风险问题，知道了理论界一个世纪以来对公共产品配置的研究、20 世纪 20～30 年代爆发的米塞斯、哈耶克与兰格、勒纳关于计划与市场的大论战，并逐步接触到新制度经济学中的产权、制度、契约，这实际上都涉及激励兼容问题。而这些问题在当今的中国显得更为突出，兴趣和为国分忧的冲动，使我深入到了对机制设计理论的学习中。由于博弈论是机制设计理论的重要分析工具，由此我开始了博弈论的学习。而当时中国已有的介绍博弈论的书真可谓凤毛麟角。张维迎所写的《博弈论与信息经济学》成为我的第一本博弈论读物，接着我读到的第二本博弈论教科书是吉本斯（Gibbons）所著的《为经济学家的博弈论》（中文译为《博弈论基础》），并收集了这本书中给出的几乎所有重要的参考文献。一直到 1999 年，台湾大学的巫和懋教授与台湾政治大学的陈恭平教授在南开讲授博弈论和信息经济学我才算正式接受了系统的课堂熏陶，本书中的许多认识都得益于两位老师的指导，在此对巫、陈两位教授表示深深谢意。最后，感谢童颖编辑辛勤劳动，她使本书增色不少。本书的撰写是在我给经济学基地班讲授中级微观经济学的过程中开始的，因而对写作的要求就较为随意，再加上自己水平有限，因而错误在所难免，在此敬请各位老师、同学批评指正，谢谢！

F224.32
Y299

姚 国 庆

2003 年 3 月于南开园